Genomics and evolution of Campylobacter: host adaptation and the emergence of globally disseminated lineages

Mourkas, Evangelos

Award date:
2020

Awarding institution:
University of Bath

Genomics and evolution of *Campylobacter*: host adaptation and the emergence of globally disseminated lineages

Evangelos Mourkas

A thesis submitted for the degree of Doctor of Philosophy

University of Bath

Department of Biology and Biochemistry

July 2020

Table of contents

Acknowledgements

I would like to express my deepest gratitude to my supervisor Sam Sheppard for his continuous support and guidance throughout my PhD. He believed in me and gave me the freedom to pursue my own ideas. He always had time for me in his busy schedule. I have gained so much more knowledge and experience in those four years than I could ever imagine.

I would like to offer my special thanks to Guillaume Méric for giving me the guidance and the push I needed at the start of my PhD and to Michelle Wille for introducing me into the world of research. I would like to thank all my collaborators for their input and help with analysis on the work done in this thesis. Namely, big thanks to Koji Yahara for his help in Chromopainter analyses for chapter 2. Dave Kelly and Aidan Taylor for their invaluable help with the phenotype assays for chapter 3 and to Diego Cuadrado Flórez for his help and input for chapter 4. Clara Atterby and Patrik Ellström for their great help on work for chapter 5.

I would also like to thank all members of the lab (Ben Pascoe, Sion Bayliss, Leonardos Mageiros, Jessica Calland, Grant Futcher, Emily Rudolph) for their friendship and scientific encouragement. Sion for constant discussions on projects, continuous support and friendship. Ben for always being there for me for help and support. Leo for being like a brother to me. My PhD-sister Jess for all support, help, friendship and for keeping me sane when I needed it the most. I am going to miss you all.

Last but not least, I would like to thank my parents for their endless love, support and for giving me the opportunity to follow my dreams. My sister Anastasia for her love and for being my pillar of strength. Finally, I would like to thank all my friends (Thanasis, Giannis, George, Greg, Tasos, Ben, Mauro, Andre, Jose, Nar, Beth and many more) for being there for me, near or far.

List of publications

The following peer-reviewed publications correspond to work which directly contributes to this dissertation:

Mourkas, E., Taylor, J.T., Méric, G., Bayliss, C.S., Pascoe, B., Mageiros, L., Calland, K.C., Hitchings, M.D., Ridley, A., Vidal, A., Forbes, J.K., Strachan, J.C.N., Parker, T.C., Parkhill, J., Jolley, A.K., Cody, J.A., Maiden, C.J.M., Kelly, J.D., Sheppard, S.K. (2020) Agricultural intensification and the evolution of host specialism in the enteric pathogen *Campylobacter jejuni*. *Proceedings of the National Academy of Sciences* 117 (20), 11018–11028.

Mourkas, E., Flórez-Cuadrado, D., Pascoe, B., Calland, J.K., Bayliss, S.C., Mageiros, L., Méric, G., Hitchings, M.D., Quesada, A., Porrero, C., Ugarte-Ruiz, M., Gutierrez-Fernandez, J., Dominguez, L., Sheppard, S.K. (2019) Gene pool transmission of multidrug resistance among *Campylobacter* from livestock, sewage and human disease. *Environmental Microbiology* 21(12), 4597–4613.

Atterby, C., **Mourkas, E.**, Méric, G., Pascoe, B., Waldenström, J., Sheppard, S.K., Olsen, B., Järhult, J.D., Ellström, P. (2018) The potential of isolation source to predict colonization in avian hosts: a case study in *Campylobacter jejuni* strains from three bird species. *Frontiers in Microbiology* 9:591.

The following peer-reviewed publications correspond to collaborations on work which did not directly contribute to this dissertation:

Jehanne, Q., Pascoe, B., Bénéjat, L., Ducournau, A., Buissonnière, A., **Mourkas, E.**, Mégraud, F., Bessède, E., Sheppard, S.K., Lehours, P. (2020) Determination of *Campylobacter coli* host-segregating SNPs and their use for source attribution of clinical isolates. *Applied and Environmental Microbiology* (Accepted).

Muñoz-Ramirez, Z.Y., Pascoe, B., Mendez-Tenorio, A., **Mourkas, E.**, Sandoval-Motta, S., Perez-Perez, G., Morgan, D.R., Dominguez R.L., Ortiz-Princz, D., Cavazza, M.E., Rocha, G., Queiroz, D.M.M., Catalano, M., de la Palma, G.Z., Goldman, C.G., Venegas, A., Alarcon, T., Oleastro, M., Vale, F.F., Goodman, K.J., Torres, R.C., Berthenet, B., Hitchings, M.D., Blaser, M.J., Sheppard, S.K., Thorell,

K. & Torres J. (2020) A five-hundred year tale of co-evolution, adaptation and virulence: *Helicobacter pylori* in the Americas. *The ISME Journal* (Accepted).

Pascoe, B., Schiaffino, F., Murray, S., Méric, G., Bayliss, S.C., Hitchings, M.D., **Mourkas, E.**, Calland, J.K., Burga, R., Peñataro-Yori, P., Jolley, K.A., Cooper, K.K., Parker, C.T., Paredes-Olortegui, M., Kosek, M. & Sheppard, S.K. (2020) Genomic epidemiology of *Campylobacter jejuni* associated with asymptomatic pediatric infection In the Peruvian Amazon. *PLOS Neglected Tropical Diseases* (in press). (medRxiv: doi.org/10.1101/2020.04.26.20075689).

Banaszkiewicz, S., Calland, J.K., **Mourkas, E.**, Sheppard, S.K., Pascoe, B. & Bania, J. (2019) Genetic diversity of composite enterotoxigenic *Staphylococcus epidermidis* pathogenicity islands. *Genome Biology and Evolution* 11 (12) 3498–3509.

Méric, G., Mageiros, L., Pascoe, B., **Mourkas, E.**, Lamble, S., Bowden, R., Jolley, K.A., Raymond, B., Sheppard, S.K. (2018) Lineage-specific plasmid acquisition and the evolution of specialized pathogens in *Bacillus thuringiensis* and the *Bacillus cereus* group. *Molecular Ecology* 27 (7), 1524-1540.

Méric, G., McNally, A., Pessia, A., **Mourkas, E.**, Pascoe, B., Mageiros, L., Vehkala, M., Corander, J., Sheppard, S.K. (2018) Convergent amino acid signatures in paraphyletic *Campylobacter jejuni* sub-populations suggest human niche tropism. *Genome Biology and Evolution* 10 (3), 763-774.

Rossi, G., Silva, H.O., Aguilar, E.G., Pascoe, B., Méric, G., **Mourkas, E.**, Hitchings, M.D., Mathias, L.A., Ruiz, V.L.A., Fukumasu, H., Sheppard, S.K., Vical, A.M.C. (2018) Comparative genomic survey of *Bacillus cereus* sensu stricto isolates from the dairy production chain in Brazil. *FEMS Microbiology Letters* 365 (3).

Eriksson, P., **Mourkas, E.**, González-Acuna, D., Olsen, B., Ellström, P. (2017) Evaluation and optimization of microbial DNA extraction from fecal samples of wild Antarctic bird species. *Infection Ecology & Epidemiology* 7 (1), 1386536.

Declarations

I can confirm that the work presented in this dissertation is the work of the author. Due to the highly collaborative nature of science, some analyses were performed by collaborators. Where this is the case, it is clearly denoted in square brackets within the text, as follows: [N.B. This analysis was performed by or in collaboration with XXXX (full name), XXXX (name) university, XXXX (city), XXXX (country)].

Chapters 2, 3, 4 and 5 are written in the style of peer-reviewed publications (chapter 2 is a draft manuscript, while chapters 3, 4 and 5 are published). References for chapters 2, 3, 4 and 5 are placed at the end of each relevant chapter. References for chapters 1 and 6, corresponding to introduction and discussion of the dissertation are placed at the end of this thesis. Supplementary figures and tables for all chapters are provided in the Appendix.

Copyright agreements and data access statements are provided for chapters 3, 4 and 5 that contain peer-reviewed publication material at the end of each chapter. Individual statements of authorship are provided at the appendix for each of the above chapters.

The material presented here for examination for the award of a higher degree by research has been incorporated into submission for another degree. Specifically, chapters 4 and 5 have been conducted in collaboration with researchers at Universidad Complutense Madrid and Uppsala University, respectively. For chapter 4, parts of work have been submitted for the degree of Doctor of Philosophy, awarded to Diego Flórez Cuadrado by the Universidad Complutense Madrid, Madrid, Spain. Specifically, the *in vitro* antibiotic resistance profiling of the isolates, along with initial analyses of the population structure, interrogation of genomes for the presence of antimicrobial resistance genes and localization of those genes in the genomes. These parts of work are clearly denoted in this dissertation in square brackets within the text, as follows: [N.B. This analysis was performed by or in collaboration with Diego Flórez Cuadrado, Universidad Complutense Madrid, Madrid, Spain]. For chapter 5, parts of work have been submitted for the degree of master's in medical sciences, awarded to Evangelos Mourkas (author of this dissertation) by Uppsala University. Specifically, the second round of infection experiments (experiment 2) and the downstream microbiological and molecular

analysis (bacteria isolation and quantification, development of real-time PCR and statistical analysis) following those experiments. These parts of work are clearly denoted in this dissertation in square brackets within the text, as follows: [N.B. This analysis was performed by Evangelos Mourkas, Uppsala University, Uppsala, Sweden].

Abbreviations and acronyms

AMR – Antimicrobial Resistance

BEAST – Bayesian Evolutionary Analysis by Sampling Trees

B. anthracis – Bacillus anthracis

CARD – Comprehensive Antibiotic Resistance Database

CC(s) – Clonal Complex(es)

cgMLST – core genome MLST

CFU – Colony Forming Units

COG – Cluster of Orthologous Groups

CI – Consistency Index

C. coli – Campylobacter coli

C. concisus – Campylobacter concisus

C. fetus – Campylobacter fetus

C. gracilis – Campylobacter gracilis

C. jejuni – Campylobacter jejuni

C. lari – Campylobacter lari

C. rectus – Campylobacter rectus

C. showae – Campylobacter showae

C. ureolyticus – Campylobacter ureolyticus

dpi – days post infection

ECDC – European Centre for Disease Prevention and Control

EFSA – European Food Safety Authority

ESS – Effective Sample Size

EUCAST – European Committee of Antimicrobial Susceptibility Testing

EU – European Union

E. coli – Escherichia coli

GB – Gene Block

GWAS – Genome-Wide Association Studies

GI – Gastrointestinal

GIs – Genomic Islands

GBS – Guillain-Barré Syndrome

GSS – Generalized Stepping-Stone Sampling

HGT – Horizontal Gene Transfer

HPD – Higher Posterior Density

IBD – Inflammatory Bowel Disease

IBS – Irritable Bowel Syndrome

KPC – *Klebsiella pneumoniae* carbapenemase

K. pneumoniae – Klebsiella pneumoniae

LMICs – Low- Middle-Income Countries

ORF(s) – Open Reading Frame(s)

MALDI-TOF – Matrix-Assisted Laser Desorption/Ionization-Time Of Flight

MS – Mass Spectrometry

MCMC – Markov Chain Monte Carlo

ML – Maximum Likelihood

mCCDA – modified Charcoal-Cefoperazone-Deoxycholate Agar

MDR – Multidrug Resistance

MLST – Multilocus Sequence Typing System

MGD – Molybdopterin-guanine dinucleotide

M. tuberculosis – Mycobacterium tuberculosis

NCBI – National Center for Biotechnology Information

NDM-1 – new Delhi metallo-β-lactamase-resistance gene

OD – Optical Density

PCR – Polymerase Chain Reaction

PBS – Peptone Buffered Saline

PFGE – Pulsed-Field Gel Electrophoresis

rMLST – ribosomal MLST

rRNA – ribosomal RNA

r/*m* – recombination to mutation ratio

R/θ – relative rate of recombination to mutation

SNP(s) – Single Nucleotide Polymorphisms

ST – Sequence Type

S. enterica – Salmonella entérica

S. typhi – Salmonella enterica serovar Typhi

S. pneumoniae – Streptococcus pneumoniae

S. suis – Streptococcus suis

tMRCA – time to Most Recent Common Ancestor

US – United States

V. cholerae – Vibrio cholerae

V. parahaemolyticus – Vibrio parahaemolyticus

wgMLST – whole genome MLST

WGS – Whole Genome Sequencing

WHO – World Health Organization

v – mean divergence of imported DNA

$1/\delta$ – inverse mean DNA import length

Abstract

Bacteria have been evolving and adapting in different hosts and environments for millions of years. Most human infections are zoonotic, occurring mainly in animals but also being transmissible to humans. Humans and animals have been living together for thousands of years and been sharing or retaining their own bacterial populations. Bacteria can adapt in new hosts and respond to different selective pressures but not all bacterial species are found in all hosts and environments. The genetic mechanisms that promote these adaptations are not fully understood. The work presented in this thesis investigates the genomic and phenotypic adaptations that promote colonization/proliferation of bacteria of the genus *Campylobacter* and explores variation at the species, lineage and gene level. Pangenomic comparative analyses revealed core and accessory gene variation highlighting the importance of gene gain and loss in the evolution of this species and the use of genomics in identifying molecular markers to monitor lineage specific *in vivo* infection experiments. *Campylobacter* are highly recombinogenic, thus a focus has been given on quantifying recombination in the genome. A detailed analysis of recombination has shown the proportion of the mobile genetic elements (mobilome) in the *Campylobacter* genus and pinpointed genes associated with host adaptation. Additionally, analysis of *Campylobacter* resistomes between species, lineages, hosts and environments revealed multidrug resistant (MDR) genomic islands (GIs) and the involvement of plasmids in horizontal gene transfer (HGT). This work has provided evidence of interspecies recombination between different species that share the same hosts and the genes associated with them. The work in this thesis has broadened understanding of how genomic plasticity can allow these versatile bacterial pathogens to adapt into new niches and environments.

Chapter 1

Introduction

Introduction

Bacteria have been on the planet for billions of years. Technological advances have allowed us to isolate and sequence thousands of bacterial genomes the past 25 years (Loman and Pallen, 2015). This has given new insights on how bacterial species are shaped and evolve. The work presented in this thesis takes advantage of large-scale sequencing data and bacterial isolates to answer questions related to species barriers, host adaptation, host ecology and antimicrobial resistance (AMR) in the *Campylobacter* genus. In this thesis, each chapter has their own focused introduction, so the following introduction will describe some general principals of how bacteria evolve, adapt to different hosts and environments and mechanisms of recombination and gene gain and loss. It will also cover how advanced sophisticated software and bioinformatics pipelines have helped us to understand how bacterial species are maintained and how anthropogenic impacts on the planet have shaped bacterial evolution and adaptation in relation to pathogen emergence and AMR.

Bacteria – a short overview

Bacteria are single cell organisms classified under the domain of Prokaryota. They can vary widely in size, but they are typically between 0.2 and 2 µm, and have different morphologies such as rod, sphere or spiral. Unlike eukaryotes, bacteria do not have a nucleus and other membrane-bound organelles. Most bacterial species have a single DNA chromosome, but there are species that have multiple circular and/or linear chromosomes (Egan et al., 2005).

Bacteria are among the most abundant organisms on the planet. Recent studies have shown that Earth is home to approximately one trillion microbial species (Locey and Lennon, 2016). In terms of biodiversity, this is a big proportion of all living organisms. Bacteria were typically identified by culturing them and characterizing their biochemical properties (Drews, 2000). However, some bacteria look alike under the microscope or have similar biochemical properties while many others are difficult to culture. This leaves a large gap in our knowledge about the true biodiversity and number of different species.

Bacteria are ubiquitous in the environment and constitute a big part of the human and animal microbiome. Around 10^{14} bacterial cells live in a single human (Whitman et al., 1998) while animals amount to around 10^{24} bacterial cells, with cattle as the

main contributor (Flemming and Wuertz, 2019). Therefore, they constitute a major component of the body. Bacteria in the gastrointestinal tract are involved in food digestion, while those found on the skin act as the first line of defence against harmful ones. Most bacteria are beneficial to the host and do not cause disease, however, there is a proportion of them that are pathogenic to humans, animals and plants. Under different conditions, these bacteria can cause mild to severe disease and, in some cases, even death. Pathogenic bacteria have evolved to colonize and adapt into new hosts, evade the immune system and attack different types of cells. These pathogenic bacteria have been extensively studied on the genetic and molecular level over the years, leading to characterization of virulence mechanisms.

Bacteria of the genus *Campylobacter*

Campylobacter jejuni (*C. jejuni*) is the leading bacterial cause of gastroenteritis in high- but also in low- and middle-income countries (LMICs), where the incidence is much higher compared to the United States (US) or the European Union (EU) (Amour et al., 2016; Pascoe et al., 2020). In addition to this, more emerging *Campylobacter* species of clinical importance have been described including *Campylobacter coli* (*C. coli*), *Campylobacter concisus* (*C. concisus*), *Campylobacter lari* (*C. lari*) and *Campylobacter ureolyticus* (*C. ureolyticus*) (Man, 2011). Bacteria of the genus *Campylobacter* are Gram-negative spiral or rod-shaped, equipped with either a single polar flagellum, a bipolar flagellar or no flagellum at all in some species (Man, 2011). They are between 0.5 to 5 μm in size and they grow under microaerobic conditions (Kaakoush et al., 2015). The first isolation of a *Campylobacter* bacterium was in the 19th century from a case of abortion in sheep (McFadyean and Stockman, 1913). The bacterium was originally named *Vibrio fetus* but was renamed as *Campylobacter fetus* (*C. fetus*) in 1963 (Sebald and Veron, 1963). The taxonomy now includes the *Campylobacter* genus in the family of *Campylobacteraceae*, the order *Campylobacterales*, the class *Epsilonproteobacteria* and the phylum *Proteobacteria*.

***Campylobacter* clinical gastroenteritis**

The two most common species known to cause human gastroenteritis are *C. jejuni* and *C. coli* (Eurosurveillance editorial team, 2014). Infection by these two species can result in watery or bloody diarrhoea, fever and stomach cramps that are usually resolved after 5 days (Man, 2011). While experimental studies in humans have

shown that the lowest infectious dose for *C. jejuni* is 800 colony forming unit (CFU) (Black et al., 1988), it is estimated that the actual infectious dose can be as low as 360 CFUs (Hara-Kudo and Takatori, 2011). The severity of the disease is not only affected by the infectious dose and the host immune's response but also by the strain itself, with different strains causing different disease outcomes (Nielsen et al., 2010). Immunocompromised patients are particularly vulnerable to campylobacteriosis than healthy individuals (Acheson and Allos, 2001). While the major cause of *Campylobacter* infections is gastroenteritis, other manifestations have been associated with this pathogen usually including conditions of the gastrointestinal tract like: inflammatory bowel disease (IBD) and postinfectious irritable bowel syndrome (IBS) (Schwille-Kiuntke et al., 2011; Gradel et al., 2009). IBDs are chronic inflammatory disorders of the gastrointestinal tract including Crohn's disease and ulcerative colitis. *C. jejuni* infections have been associated with an increased risk of developing IBD (Gradel et al., 2009). More recent studies demonstrated evidence of the association of other *Campylobacter* species, like *C. concisus*, *C. ureolyticus*, *Campylobacter showae* (*C. showae*), *Campylobacter rectus* (*C. rectus*) and *Campylobacter gracilis* (*C. gracilis*), with IBD (Man et al., 2010; Zhang et al., 2009). Furthermore, patients that had a severe enteritis caused by *Campylobacter* have been associated with developing a subsequent postinfectious irritable bowel syndrome (Schwille-Kiuntke et al., 2011).

Campylobacter - extraintestinal manifestations

Extraintestinal manifestations have also been associated with *Campylobacter* infections following gastroenteritis. These include Guillain-Barré syndrome (GBS), reactive arthritis, bacteraemia and sepsis. GBS is a neurologic disorder and is caused by the human body's immune system attacking the nervous system (van Doorn et al., 2008). Outbreaks of GBS have occurred following outbreaks of *C. jejuni* infections (Jackson et al., 2014). The syndrome usually appears postinfection, following *C. jejuni* gastroenteritis, in 1.2 to 2.3 cases per 100,000 cases annually (van Doorn et al., 2008). Human antibodies, produced as a response to *C. jejuni* lipooligosaccharides, cross-react with the human gangliosides, leading to weakness and tingling in limbs which can progress to the respiratory system, potentially leading to paralysis and death (Freddo et al., 1986). Reactive arthritis usually affecting joints has also been described in cases following a gastroenteritis caused by *C. jejuni*. The symptoms often appear a month postinfection and are resolved

after a year, though, in some cases, they can persist up to five years (Batz et al., 2013). Bacteraemia and septicaemia have been reported in cases of immunocompromised patients with underlying health problems (Man, 2011). This has been associated with several *Campylobacter* species, including *C. jejuni*, *C. coli*, *C. fetus* and *Campylobacter upsaliensis* (*C. upsaliensis*) (Man, 2011; Chusid et al., 1990).

Campylobacter – pathogenesis

The pathogenesis and the mechanisms underlying the cause of disease are poorly understood in *Campylobacter*, mainly due to the lack of a good animal model (Young et al., 2007). Murine models fail to replicate the human disease as well as having inconsistent results (Fox et al., 2004; Mansfield et al., 2007). Ferrets are a good animal model that develop symptoms of the disease that are seen in humans (Fox et al., 1987), but the cost of experiments is high (Young et al., 2007). Chickens are a natural host for *Campylobacter* and a good animal model to study the colonization factors, as well as a tool to design anti-*Campylobacter* strategies and interventions to prevent spillover to human populations (Young et al., 2007). Comparative genomic analyses have revealed mechanisms of pathogenicity and chicken colonization factors. Pathogenicity in *Campylobacter* appears to be multi-factorial with the flagella, capsule, O- and N-linked protein glycosylation systems and secreted proteins being implicated with host-cell invasion (Parkhill et al., 2000; Karlyshev et al., 2002; Gilbert et al., 2002). The flagellum is involved in colonization, virulence, epithelial cell adherence and invasion (Young et al., 2007). Lipooligosaccharides (LOS) and the capsule are involved in immune avoidance (Nachamkin, 2002). The genes that encode for the flagellum, capsule and LOS are found in highly variable regions of the genome, where mutations have been associated with variability in serum resistance, cell adherence and invasion of human intestinal epithelial cell lines (Fry et al., 2000). To establish infection, *Campylobacter* bypasses the mucus layer of the GI epithelium owing to their motility and unique corkscrew morphology (McSweegan and Walker, 1986). Once they are inside they interact with the underlying epithelial cells (Biswas, 2000; Monteville, 2003). Infection experiments in ferrets have shown a role of type IV secretion system in cell invasion and pathogenicity (Bacon et al., 2002). The immune response, following intestinal epithelial cell invasion, involves cytokine induction with IL-8, a proinflammatory cytokine, commonly stimulated in human infection (Hickey et al.,

2000). *Campylobacter* causes diseases in humans but behaves as a commensal in chickens indicating host-specificity as a reason behind differences in immune responses (Young et al., 2007).

Campylobacter - epidemiology

Epidemiological data from the last 15 years suggest an increasing trend in campylobacteriosis cases worldwide (Kaakoush et al., 2015). However, there are differences in numbers between countries and regions, which may reflect differences in sampling and isolation methodologies but also surveillance bias and underreporting in hospitals (Kaakoush et al., 2015). In LMICs, the exposure rates are much higher (Martin et al., 1988), and the disease appears to be endemic especially among young children, usually as a chronic recurrent infection (Coker et al., 2002). Asymptomatic carriage appears to be very common among children in LMICs (Lee et al., 2013), but the exact mechanisms underlying the variation in disease symptoms are not well understood (Pascoe et al., 2020). In the US the reported annual incidence of *Campylobacter* infections was recorded by the Food-Borne Diseases Active Surveillance Network as 14.3 per 100,000 population for the period between 1996 and 2012 (Crim et al., 2015). The total number of campylobacteriosis reported cases in EU is around nine million, surpassing cases of salmonellosis (Havelaar et al., 2013). It is estimated that the true incidence is eight to 30 times higher than the confirmed cases due to underreporting and the self-limiting nature of the disease (Janssen et al., 2008).

Campylobacter - isolation methodologies

Isolation methods and protocols for *C. jejuni* and *C. coli* have been developed for samples from clinical patients and poultry meat (Chon et al., 2013). Modified charcoal-cefoperazone-deoxycholate agar (mCCDA) along with other selective media has been used to isolate *Campylobacter* bacteria (Corry et al., 1995; Oyarzabal et al., 2005). Selective media for *Campylobacter* contain several antibiotics to suppress the growth of competing bacterial flora in samples (Corry et al., 1995). The most common antibiotics used are cefoperazone and vancomycin, against which *Campylobacter* is resistant (Corry et al., 1995; Oyarzabal et al., 2005). However, resistance to these antibiotics has spread to other bacteria including *Escherichia coli* (*E. coli*), making the detection and differentiation of *Campylobacter* bacterial colonies a difficult task (Jasson et al., 2009; Moran et al., 2009). *C. jejuni*

and *C. coli* are highly thermotolerant, growing at temperatures between 37 ºC and 42 ºC, suppressing the growth of other bacterial species (On, 2013). Enrichment methods for isolation of *Campylobacter* in samples that have a low number of cells have been reported including incubation in various broths (Bolton broth, brucella broth, Preston broth) (Kim et al., 2009) and coculture with amoebas (Axelsson-Olsson et al., 2007).

Laboratory diagnosis of clinical *Campylobacter* infection is based primarily on culture-dependent methodologies where single colonies are isolated and subjected to a range of biochemical tests for species identification (Kaakoush et al., 2015). Diagnosis is also performed with culture-independent techniques which require the isolation of DNA or RNA from the clinical sample or from a pure culture. Different genetic markers can be used to identify the genus and/or the species by polymerase chain reaction (PCR) amplification of a gene or genes of interest. The 16S ribosomal RNA (rRNA) gene is commonly used for detection and differentiation of *Campylobacter* species causing enteritis (Kulkarni et al., 2002). A recent tool for rapid bacterial species identification in microbiological diagnostics, is the matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectrometry (MS) (Wieser et al., 2012). This method can identify, with high reliability and accuracy, the bacterial species directly from colonies from a pure culture making it a valuable tool in the hospital's microbiological diagnostics facilities (Wieser et al., 2012). Whole genome sequencing (WGS) is now very common for bacterial species identification and is becoming part of hospital diagnostics protocols.

Campylobacter - risk factors, ecology, transmission and environmental reservoirs

The main risk factor for campylobacteriosis is international travel (Swaminathan et al., 2009), followed by consumption of contaminated meat, particularly poultry products (Hermans et al., 2012; Fravalo et al., 2009; Guyard-Nicodème et al., 2013). Other risk factors include consumption of unpasteurized milk or untreated water (Levesque et al., 2008), handling, preparing of contaminated meat during slaughter, including poultry, pork and beef (Fravalo et al., 2009; Hermans et al., 2012; Guyard-Nicodème et al., 2013). *C. jejuni* is one of the six most common pathogens associated with travelling abroad, with the travel destination playing an important role in the risk of getting infected (Swaminathan et al., 2009). South and southeast

Asia, Latin America and Africa are shown to have a high risk of infection with *C. jejuni* (Mughini-Gras et al., 2014). In the EU, 12% of all campylobacteriosis cases in 2010 were linked to international travellers (Gautret et al., 2012). Quantifying the contribution of different infection sources to human campylobacteriosis cases is of importance to public health. An approach to this was introduced by Maiden et al in 1998 using molecular typing methods including the multilocus sequence typing system (MLST) (Maiden et al., 1998).

Attributing clinical *C. jejuni* and *C. coli* infections to existing genotypes commonly isolated from different animal hosts revealed a strong link with ecology (Sheppard et al., 2011). Poultry is a main source of getting *Campylobacter* with broiler chickens having high carriage of *C. jejuni* and *C. coli* (Skirrow, 1977; Kaakoush et al., 2015; Sheppard et al., 2009b). Furthermore, farm and surrounding water environments are also contaminated with *C. jejuni* due to spillover from poultry faeces (Figure 1.1) (Ellis-Iversen et al., 2012). Other farmed birds like turkeys and ducks are also known to carry these two *Campylobacter* species (Giacomelli et al., 2012; Colles et al., 2011). Source attribution studies based on MLST data from >2000 studies worldwide estimate that the majority of human campylobacteriosis cases come from chickens (Cody et al., 2019). Recent studies of the UK Food Standards Agency have showed that around 70% of fresh retail chickens sampled, had high levels of *Campylobacter* (Agency, 2014). Cattle also contribute to campylobacteriosis cases globally with around half a million cases in EU caused by two of the most common cattle associated *C. jejuni* lineages (Mourkas et al., 2020). Eating habits and cultures might affect the sources of infection in different countries and regions, with the number of cases attributed to cattle, equally high as the ones to chicken in France (Thépault et al., 2017). Investigating the source of human campylobacteriosis infections and assigning it to different host reservoirs based on gene and allelic similarities is valuable to design intervention strategies.

Livestock including chickens, cattle, pigs, sheep, goats also carry these bacteria in their gastrointestinal tract and this is a major reservoir of *Campylobacter* species (Figure 1.1) (Sproston et al., 2011; Sheppard et al., 2009a; Ogden et al., 2009). Companion animals and pets are also found to be carriers with dogs, cats, rabbits as well as pet reptiles having tested positive for *Campylobacter* (Figure 1.1) (Ogden et al., 2009; Giacomelli and Piccirillo, 2014; Kohler et al., 2008). Wild animals,

specifically wild birds, are usually carriers of *Campylobacter* species (Figure 1.1). Various studies on different wild bird species have shown the abundance of *Campylobacter* species, including *C. jejuni*, *C. coli* and *C. lari* (Sheppard et al., 2009a; Sheppard et al., 2011; Griekspoor et al., 2013; Griekspoor et al., 2009; Sproston et al., 2011; Leatherbarrow et al., 2007). Water contaminated with *Campylobacter* has been responsible for various outbreaks and together with wild birds play a role in transmission to humans (Figure 1.1) (Hänninen et al., 2003; Rogol et al., 1983). Wastewater and private water supplies can be contaminated with *Campylobacter*, indicating possible sources of transmission to humans and animals (Ellis-Iversen et al., 2009). Being a thermotolerant microaerobic bacterium, it is not clear how *Campylobacter* survives in the water, but studies suggest a role for amoebas as potential reservoirs for this pathogen in the environment (Axelsson-Olsson et al., 2005). *C. jejuni* was found to escape digestive vacuoles inside the amoebas, multiply and subsequently rupture the amoebic cell (Axelsson-Olsson et al., 2005). Many microbial organisms have been detected in farm water systems (Snelling et al., 2006), suggesting a role in transmission to livestock.

Campylobacter **- treatment and antibiotic resistance**

Most human cases of campylobacteriosis are self-limiting with the infection resolving usually after 3-5 days. However, in severe cases or in immunocompromised patients, antibiotic treatment may be required, with fluoroquinolones and macrolides being the drugs of choice (Acheson and Allos, 2001). High levels of AMR in *C. jejuni* and *C. coli* against ciprofloxacin, a fluoroquinolone widely used for campylobacteriosis treatment, has become a worrying problem (Food and Authority, 2019). It has been demonstrated that a single point mutation in the *gyrA* gene can confer resistance to fluoroquinolones (Luo et al., 2003; Gibreel, 2006). This has led to macrolides as an alternative drug option, and more specifically erythromycin, where AMR develops relatively slowly with point mutations arising in the 23S rRNA gene (Lapierre et al., 2016). However, the detection of an rRNA methylating enzyme, encoded by the *ermB* gene conferring resistance against erythromycin was recently discovered in China (Qin et al., 2014) and later in EU (Florez-Cuadrado et al., 2016; Florez-Cuadrado et al., 2017). More worryingly, an increasing trend of MDR *Campylobacter* isolates, specifically *C. coli*, has been observed the last 10 years with many isolates carrying multiple AMR genes conferring resistance against several drug classes (Luangtongkum et al., 2009; Food and Authority, 2019).

Antibiotics have extensively been used as growth promoters in animal feed all over the world (banned in EU in 2006) (Castanon, 2007). This has led to development of AMR in bacteria of the gastrointestinal tract of farmed animals (Sproston et al., 2018; Sheppard et al., 2009a; Sheppard et al., 2009b). Close contact between agricultural animals and humans is increasing the risk of transmission of AMR bacteria with studies highlighting the likelihood of reciprocal spread of resistance (Boerlin and Reid-Smith, 2008; Huttner et al., 2013).

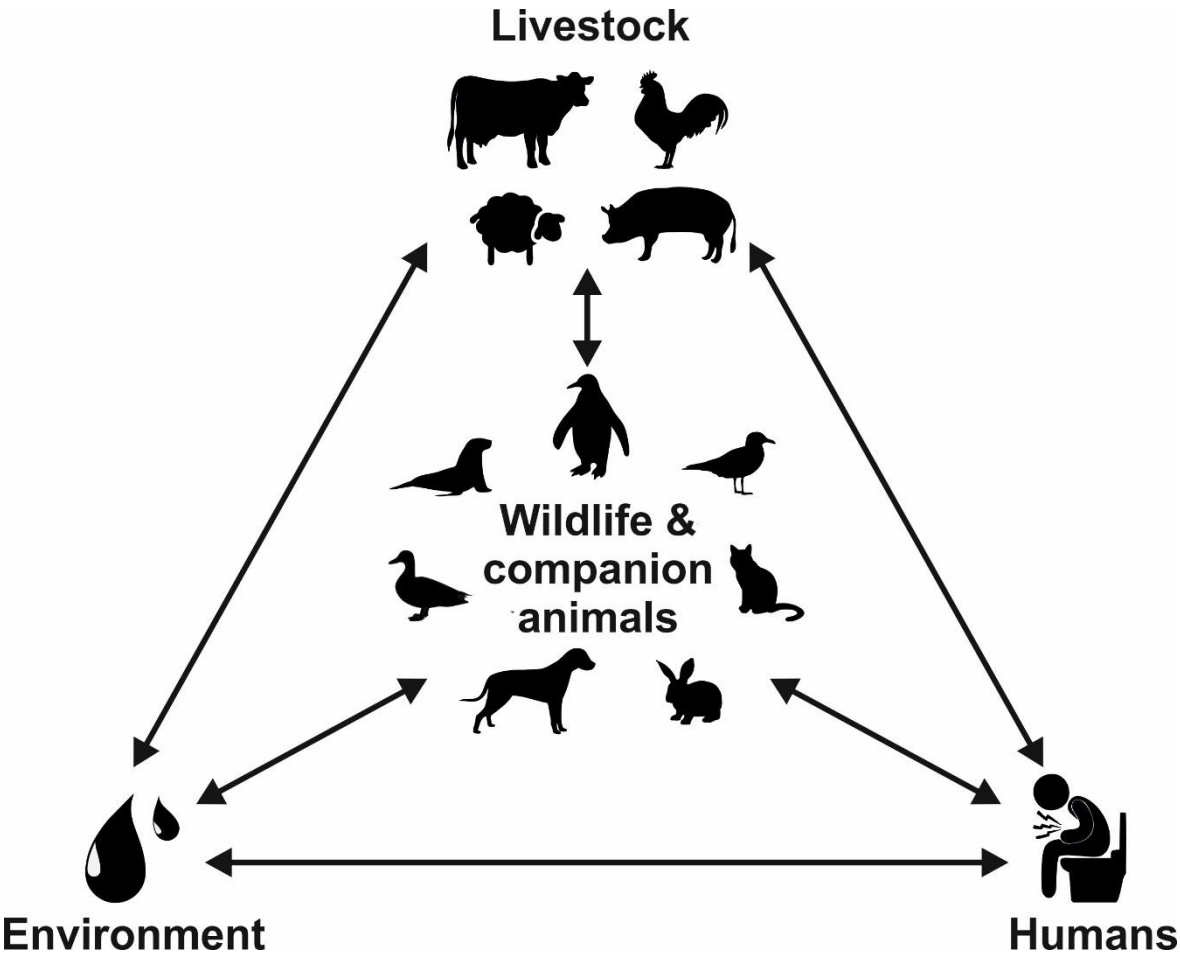


Figure 1.1. Host range and transmission of *C. jejuni*. The wildlife reservoir comprises of wild birds and marine mammals (known spillover to seals). The livestock reservoir comprises chickens, ruminants and pigs. Solid lines represent routes of transmission of *C. jejuni* between hosts and the environment.

Bacterial typing methods

Bacteria live and replicate in various ecological reservoirs including animals, food products, plants, water, humans and the environment. Clusters of potentially

pathogenic bacteria from these sources can transmit and infect human populations. These infections are usually described as outbreaks and can lead to dangerous epidemics unless the causative agent is identified, and the outbreak controlled. Identifying the genotype of bacterial isolates or their phenotypic characteristics is known as bacterial typing (van Belkum et al., 2007). One of the first typing methods, still routinely used in many parts of the world today, is Pulsed-Field Gel Electrophoresis (PFGE) (Nsofor, 2016). PFGE separates DNA fragments on agarose gel based on their size. The genomic DNA is fragmented using restriction enzymes and run on a gel to check for variation between isolates (Macfarlane et al., 1999). This method is quick and enables for identifying variation among isolates, however it does not work well with large DNA fragments. A serotyping system, used for the past 30 years for phenotypic discrimination of *Campylobacter* spp, is the Penner serotyping scheme or Heat Stable serotyping system. This system was based on a passive slide hemagglutination assay using soluble heat stable antigens (Penner and Hennessy, 1980). Many other typing systems have been developed to describe strain variation within bacterial species, something particularly helpful in outbreak investigations to discriminate cases from the background populations. Such systems, like the Arbitrarily Primed Polymerase Chain Reaction (Roberts et al., 1998) and the Multi Locus Enzyme Electrophoresis (Selander et al., 1986; Milkman, 1973) were useful in looking for variation in bacterial isolates before the advent of more precise sequence-based approaches.

However, the above-mentioned techniques have several limitations. First, they can be slow and laborious, and second, they can be hard to standardize, making it difficult to compare the results among different laboratories. In addition, early typing methods were largely designed to address species and were not generalizable across genera. A breakthrough in bacterial typing arrived with the development of sequence-based approaches that focus on variation in specific loci. In the 1990's, Profs Maiden and Spratt at Oxford University, and Prof Achtman at the Max Plank Institute in Berlin, developed MLST to characterize different species (Maiden et al., 1998; Achtman, 1996; Enright and Spratt, 1999; Enright and Spratt, 1998). MLST has been successfully used for different bacterial species like *Campylobacter* (Dingle et al., 2001), *Streptococcus* (Enright and Spratt, 1998) and *Helicobacter* (Suerbaum et al., 1998). Using this approach, the genetic variation is studied by comparing specific genes instead of variable sequence sites in the genome. These

techniques target specific housekeeping genes, usually seven, encoding enzymes with important metabolic functions for the bacterial cells. For the *Campylobacter* MLST scheme these genes are: *aspA* (aspartase A), *glnA* (glutamine synthetase), *gltA* (citrate synthase), *glyA* (serine hydroxymethyltransferase), *pgm* (phosphoglucomutase), *tkt* (transketolase) and *uncA* (ATP synthase A subunit) (Dingle et al., 2001). This typing system assigns each gene a unique arbitrary allele number based on sequence variation at that locus. Each isolate is then assigned seven numbers that classify its sequence type (ST). STs can be further classified into clonal complexes (ST-complex) if they share at least four out of the seven housekeeping genes (Dingle et al., 2001). This technique has been widely adopted and uses online databases to which sequences are uploaded and stored along with isolate metadata. Different web-based platforms are now freely available to the public for different bacterial species (http://pubmlst.org/). These databases can be used by laboratories from all over the world. At the time of writing, there were > 100,000 isolates in the *Campylobacter* pubMLST database.

The MLST system has been a great advancement in creating a standardized typing method in combination with the availability of online databases for archiving and studying thousands of isolates. However, this system accounts for genetic variation in only seven genes (<1% of the genome). The development of high-throughput next generation sequencing has revolutionized the way we think of genomes. It is now relatively inexpensive to perform WGS, costing around £50 per genome. Along with the rapid developments in software for sequence analysis, this has led to new insights in the population genomics of bacteria. However, MLST remains a useful reference as many clonal complexes represent sequence clusters that are recovered in whole genome phylogenies. For this reason, MLST clonal complex designations are still commonly used to describe observations in WGS data. In the genomic era, seven-locus MLST methods have given rise to gene-by-gene analysis approaches (Sheppard et al., 2012; Maiden et al., 2013; Thépault et al., 2017) that include increasing numbers of genes for discriminating isolates. These include ribosomal gene (rMLST) (Keith A. Jolley et al., 2012), core genome MLST (cgMLST) (Keith A Jolley et al., 2012) and whole genome MLST (wgMLST) (Jolley and Maiden, 2010; Sheppard et al., 2012) in which more genes are included in the scheme (Maiden et al., 2013).

What constitutes a bacterial pangenome?

HGT or lateral gene transfer is the movement of genetic material between organisms (not including the vertical DNA transfer from parent to offspring). Prokaryotes acquire genetic material from an external source and incorporate it into their chromosome via recombination. Comparative genomics of thousands of strains belonging to different bacterial species, brought to light major differences in genome content between species but also between strains within species. It quickly became apparent that some chromosomal regions are shared with most isolates of the same species (core genome) while others vary between different strains (accessory genome) (Figure 1.2). The combination of core and accessory genome constitutes the pangenome (Figure 1.2). Some species have been described as having open pangenomes, with new accessory genes continuously found as more strains are sequenced, while other species have closed pangenomes with few gene content differences (Figure 1.2) (McInerney et al., 2017).

By sequencing the whole genome of different bacterial species, researchers were able to define the genome size (number of bp for a single genome) and the number of genes. The genome size for different bacteria can vary widely from 112 Kb to 14 Mbp (Bennett and Moran, 2013; Han et al., 2013). Sequencing thousands of isolates for hundreds of different species showed that there is big variation between core and accessory genes for each species. Different studies have analysed the pangenomes of large-scale sequence data for various well-known bacterial species and have highlighted these differences. Analysis of >2,000 *E. coli* genomes has identified 3,188 core gene families (core genes defined as genes present in >95% of all isolates) and a total of around 90,000 unique gene families (accessory genes) (Land et al., 2015). This is a high level of variation compared to other bacteria like *C. jejuni* with 1,432 core gene families (>95%) and a total of approximately 2,422 unique gene families (accessory genes) in >1,000 isolates (Mourkas et al., 2020). For the asymptomatic opportunistic pathogen, *Staphylococcus epidermidis*, a study has identified 1,946 core genes and 12,079 unique genes present in 415 isolates (Méric et al., 2018). A study looking at 67 genomes of the intracellular pathogen *Chlamydia trachomatis* identified a total of 927 genes in the total pangenome, out of which 821 were core genes (McInerney et al., 2017).

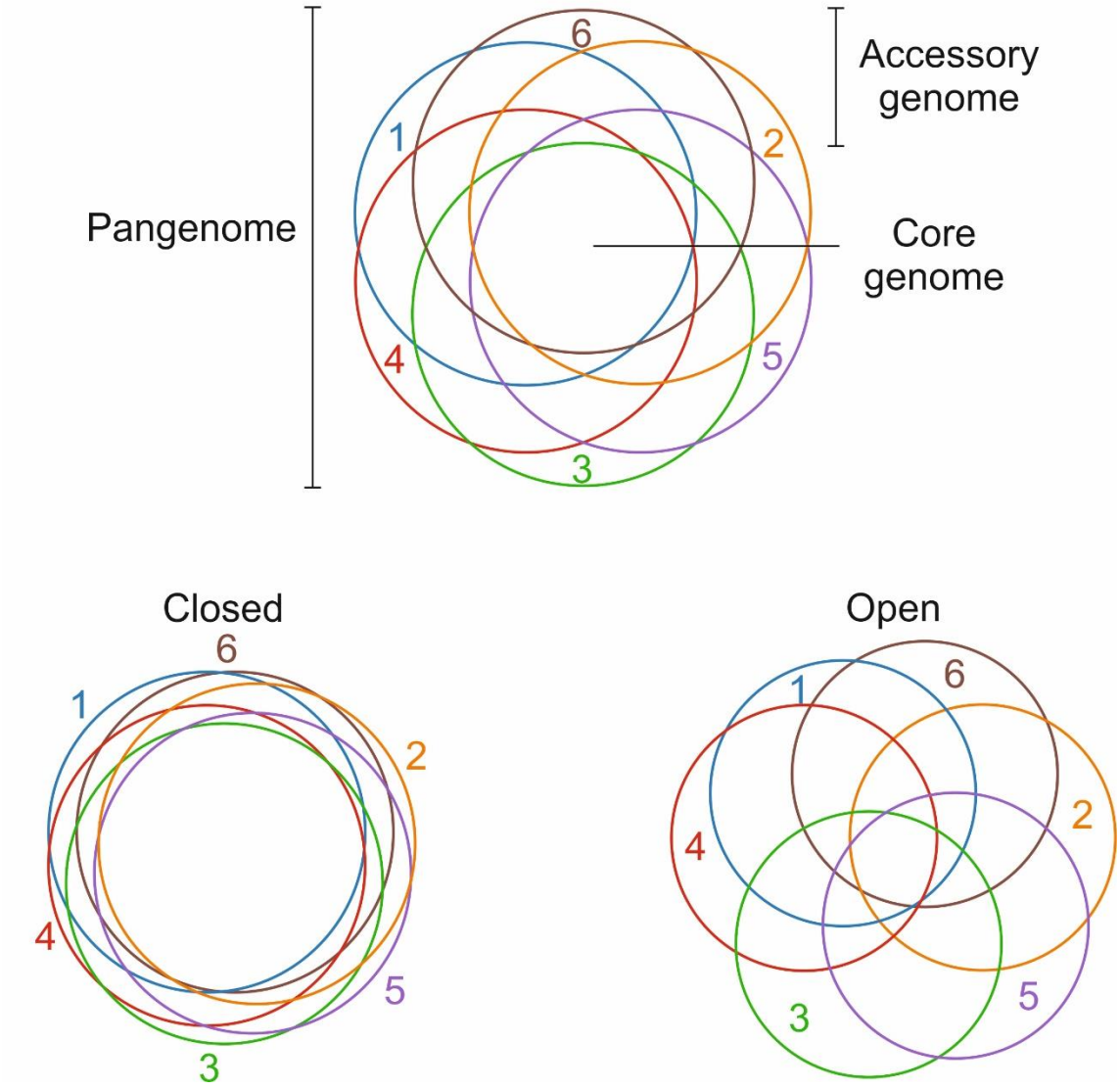


Figure 1.2. Schematic representation of pangenomes as Venn diagrams. Numbers 1 to 6 correspond to different bacterial lineages. Lineages differ in the sizes of their pangenomes, with some having closed or open pangenomes. Figure inspired and adapted by McInerney et al., 2017.

These numbers vary as more isolates of bacterial species are sequenced with the number of core genes getting slightly lower while the number of accessory genes continues to increase. Even though the number of genomes used in these studies varies significantly, there is a clear understanding that there are differences between the pangenomes of those bacterial species. A study by McInerney et al. suggests that “pangenomes are the result of adaptive, not neutral, evolution” by looking at models of HGT, the main force that shapes differences in gene content between different species (McInerney et al., 2017). They argue that when large populations with large pangenomes occupy various niches, they adapt better and that gene gain enhances this adaptation (McInerney et al., 2017). Others argue that having a larger

effective population size gives more genetic diversity and hence these species seem to have very diverse pangenomes (Andreani et al., 2017). Although these theories are not mutually exclusive it is important to consider genes as unique entities since many genes inhabit many different niches and microniches within a certain environment. Every different bacterial species has their own proportion of mobile genes (mobilome) and so an alternative way would be to apply these theories to gene, rather than species level (Shapiro, 2017). Host-associated gene pools consist of these mobile genes that are horizontally transferred between bacterial strains that are occupying a specific niche constituting the resident host microbiota (Sheppard et al., 2018).

The role of horizontal gene transfer and gene loss in bacteria

One of the main questions about accessory genes in pangenomes is the mechanisms by which these genes are transferred between bacteria. It is known that many genes arise by gene duplication and HGT (Figure 1.3). The effect of gene duplication in the genome is the generation of multiple copies of the same gene, known as paralogous genes. HGT was a known mechanism of transfer of DNA between bacteria even before the development of modern genome sequencing. There are three different forms of HGT: (i) transformation, which is the uptake of extracellular DNA, (ii) conjugation, which is the exchange of DNA between two bacterial cells and (iii) transduction, which is exchanging genetic material mediated by viruses commonly known as bacteriophages (Figure 1.3) (Mira et al., 2010).

(i) Natural bacterial transformation is the process where exogenous DNA is taken up by a recipient cell (Figure 1.3). This enables bacteria to acquire new genetic elements and potentially adapt to new environmental conditions (Thomas and Nielsen, 2005). For transformation to occur, bacteria need to be in a competent state under specific conditions (Solomon and Grossman, 1996). Bacterial transformation was first described in *Streptococcus pneumoniae* (*S. pneumoniae*), and since then, 82 more species have shown to be transformable (Thomas and Nielsen, 2005). Most bacteria like *S. pneumoniae*, *Bacillus subtilis* and *Haemophilus influenzae* regulate this competence while others, including *Neisseria gonorrhoeae*, seem to acquire genes without a regulatory system (Hamilton and Dillard, 2006).

(ii) Conjugation takes place when there is transfer of mobile genetic elements between two bacterial cells and is usually mediated by a pilus, a bridge-like connection between two cells (Figure 1.3) (Thomas and Nielsen, 2005). These genetic elements constitute conjugative plasmids and are transferred between bacteria of the same or different species (Figure 1.3). These are usually responsible for disseminating antimicrobial and virulence genes among pathogenic bacteria (Llosa et al., 2002). This process can take place between bacteria that belong to the same species, but interspecies gene transfer has been described in many bacteria like *Campylobacter* (Mourkas et al., 2019). An example of interspecies plasmid transfer has been described between Gram-positive and Gram-negative bacteria spreading the *tetO* gene which confers resistance to tetracycline (Batchelor, 2004; Taylor et al., 1983).

(iii) Transduction is the process of transfer of DNA from one bacterium to another via viruses (Figure 1.3). These viruses, known as bacteriophages or phages, infect bacteria and exploit the host cell mechanisms to replicate, by incorporating their DNA into the bacterial genome (lysogenic cycle) or by replicating inside the bacterial cell (lytic cycle) (Balcazar, 2014). Recent advances in high-throughput sequencing techniques has shown the important role of bacteriophages in the evolution of bacteria (Hatfull, 2008). This has significant implications on the spread of antibiotic resistance, as phages could transfer AMR genes between bacterial species (Balcazar, 2014). A new distinct form of transduction, called lateral transduction, was recently described in *Staphylococcus aureus* (*S. aureus*) highlighting the impact of bacteriophages in the evolution of bacterial adaptation (Chen et al., 2018).

Bacteria can acquire DNA, and thus different genes, by HGT in form of plasmids, transposons (small fragments of DNA that move around in the genome carrying genetic elements along with them) and phages. One of the main forces of bacterial evolution is the acquisition of new genes via HGT. When bacteria find themselves in new environments, HGT can introduce genes which provide novel metabolic mechanisms essential for surviving in the new niche (Lawrence, 1999). Gene gain has been described in various bacteria including *Streptococcus*, *Staphylococcus*, *Prochlorococcus* and *Campylobacter* as a way of evolving to adapt into a new host or environment (Kettler et al., 2007; Marri et al., 2006; Noto et al., 2008; Mourkas et al., 2020).

Gene loss is another important mechanism that shapes bacterial evolution, potentially altering phenotypes and promoting adaptation in a new niche. This can occur by a series of mutations following loss of function of a gene. This is particularly evident in bacterial obligate endosymbionts that have evolved from free-living bacteria with larger genomes (Toft and Andersson, 2010). These endosymbiotic bacteria have smaller genomes as a result of increased genetic drift and genetic bottlenecks (Moran and Wernegreen, 2000). The accumulation of mutations can lead to non-functional genes, called pseudogenes, by either generating truncated proteins or affecting essential amino acid positions (Land et al., 2015). In addition, insertions or deletions can lead to frameshifts that result in loss of gene expression (Land et al., 2015). Under laboratory conditions, the deletion of chromosomal positions in *Salmonella* resulted in beneficial fitness effects compared to the parental strain (Koskiniemi et al., 2012). Transposon-insertion associated mutations are associated with enhancing the growth rate of *E. coli* in nutrient-limited environments (Hottes et al., 2013; Lenski, 2017). In *Listeria* species, the loss of virulence genes was associated with change in lifestyle shifting from a facultative pathogen to saprotroph (den Bakker et al., 2010). In *C. jejuni*, the loss of a genomic island consisting of nine genes was associated with adaptive evolution following a host transition to cattle (Mourkas et al., 2020). Another example of gene loss promoting adaptive evolution is highlighted in *Mycobacterium* species following a host transition from humans to animals (Gagneux, 2018; Brosch et al., 2002). Gene loss reflects a selection against maintaining specific genes in specialized bacteria and is likely playing a major role in promoting adaptation of various bacterial species into their hosts and environments. Another effect following a bottleneck event is the accumulation of pseudogenes (non-functional segments of DNA that resemble functional genes) (Sheppard et al., 2018).

Impact of recombination in the evolution of bacteria

Recombination in bacteria takes place when there is DNA transfer between a donor and a recipient cell. The mechanism by which this can occur is by transformation, conjugation or transduction. There are two types of recombination known as homologous and non-homologous. This classification depends on how the DNA is integrated in the recipient genome. Non-homologous recombination involves the introduction of a new segment of DNA (or gene) by HGT (Figure 1.3). On the other

hand, homologous recombination involves replacement of existing DNA with new allelic variation of already present genes (Figure 1.3) (Vos, 2009). These are both potential adaptive strategies which increase the genetic variation in a population of bacteria (Otto and Lenormand, 2002). However, this might be costly for the bacteria if it disrupts already beneficial allelic combinations that have survived through many events of natural selection (Vos, 2009).

In eukaryotes, specific positions in the genome, described as hotspots, are found to be more recombinogenic (Petes, 2001). Recently, in bacteria, a population genetic method that allows for recombination rate inference, identified recombination 'hot' regions in human pathogenic isolates from ten different bacterial species (Yahara et al., 2016). Homologous recombination rates vary between different bacterial species with some of them having high rates (Vos and Didelot, 2009). While this is not well understood yet, it is likely that bacteria need DNA repair mechanisms following recombination, or they need DNA for essential metabolism purposes (Vos and Didelot, 2009). Most studies have focused on calculating the rates based on seven MLST genes (Wilson et al., 2009; Vos and Didelot, 2009). Vos et al. discussed these rates in different species and found that some appear to be very recombinogenic such as *Vibrio parahaemolyticus* and *S. pneumoniae*, others intermediate, including *C. jejuni* and *Enterococcus faecium* and others low, including *S. aureus* and *Clostridium difficile* (Vos and Didelot, 2009). However, these studies focused only on seven genes which are highly conserved between species and therefore the recombination rates are probably very different across the whole genome (Didelot et al., 2012; Croucher et al., 2013). Recent studies have shown that recombination rates are much higher and vary across different lineages in *C. jejuni* (Mourkas et al., 2020).

Homologous recombination does not necessarily alter the bacterial genome size, as one allelic variant can be replaced by another, a process known as gene conversion. This can introduce challenges when inferring relationships between isolates as two or more isolates can acquire convergent genetic changes without sharing a common ancestor as a result of genome adaptation into the same niche (convergent evolution). These isolates would appear closely related but would, in fact, be from divergent ancestry. At the same time, isolates could appear to be divergent due to recent acquisition of DNA when, in fact, they are closely related. To counteract

misinterpretation of ancestral relatedness it is important to use as much of the genome as possible when inferring relationships between isolates. Homologous recombination can skew phylogenetic relationships and thus software that account for the clonal frame have been developed to cope with this problem (Didelot and Wilson, 2015). If the recombination hotspots are removed, then the remaining genome would reflect the clonal frame (Didelot et al., 2009). ClonalFrameML can detect recombination and account for it when reconstructing the phylogeny of a multiple sequence alignment (Didelot and Wilson, 2015).

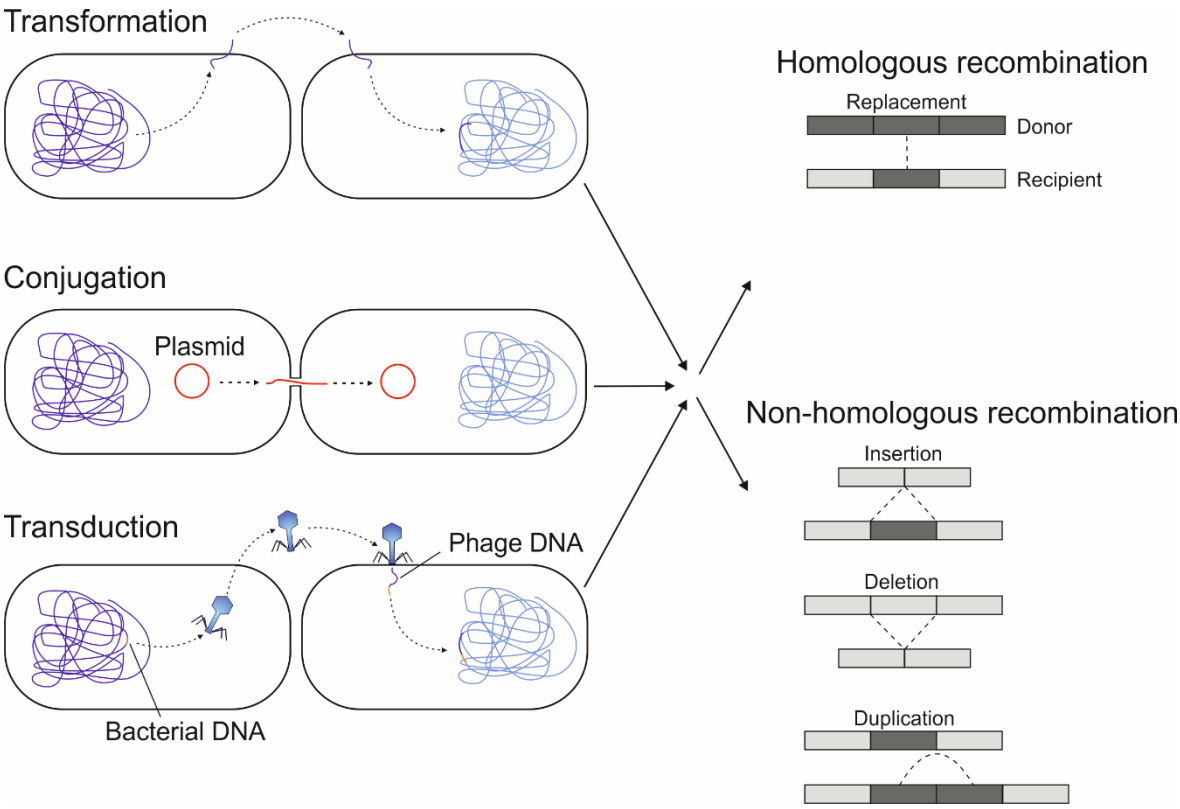


Figure 1.3. Genetic mechanisms of HGT. Transformation: the recipient bacterium takes up extracellular DNA from a donor bacterium; conjugation: exchange of DNA via plasmids between a donor and a recipient bacterium; and transduction: DNA is transmitted by a bacteriophage from a donor to a recipient cell and is integrated in the recipient cell genome along with the phage DNA. In homologous recombination DNA is replaced with a homologous sequence from another lineage. In non-homologous recombination DNA is inserted or deleted. This might lead to gene duplication. Figure inspired and adapted by Sheppard et al., 2018.

Bacterial population structure and genetics

The study of bacterial genetic diversity is closely linked with bacterial population studies. Genetic diversity is been shaped by evolutionary forces over time. These forces have led to the evolution of different bacterial species and sequence clusters of groups of isolates within species. But how do bacterial populations spread and

evolve to create the diversity and known population structure that we observe in natural populations? Bacteria reproduce by binary fission where the daughter cell is identical to the parental one, but also exchange genomic elements through recombination (Shapiro, 2016; Smith et al., 1993). During cell division, random mutations occur which are passed down to the offspring cells leading to the evolution of various clones in a bacterial population. Each of these mutations, inherited independently, would be maintained in the same position of the genome in a clonal population (linkage disequilibrium). This is, in theory, an example of perfectly clonal bacteria, but this idealized scenario is not true, as most bacterial species are affected by recombination at time points in their evolution (Shapiro, 2016). Recombination breaks down this linkage disequilibrium but, in most species, the overall genetic structure is maintained. Linkage disequilibrium is affected by the rate of mutation, genetic drift and the population structure (Smith et al., 1993).

The rate of recombination varies between bacteria and across the genome. Some species are largely clonal while others are highly recombinogenic with genetic admixture all but abolishing the clonal frame (panmictic) (Smith et al., 1993; Yahara et al., 2016). Pathogenicity islands that contain virulence or AMR genes are among known highly recombining regions and have previously been described (Hacker et al., 1997). Examples of pathogenic bacteria include *Campylobacter* (Wassenaar et al., 1995)*, Pneumococcus* (Hanage et al., 2009) and *Listeria* (den Bakker et al., 2008), highlighting the role of HGT in the mobilization of these genes (Mourkas et al., 2019). In agricultural associated bacteria, rapidly recombining genes can be linked to host ecology, metabolism and diet (Weinert et al., 2015; Richardson et al., 2018). However, in non-pathogenic bacteria these genes can be related to adaptation to a new environment (Coleman, 2006). Smith et al., analysed different species looking at chromosomal genes and found that species range from being fully clonal to panmictic (Smith et al., 1993). Some clonal species appear to have emerged from a panmictic gene pool. A typical example is the genus *Vibrio*, including *Vibrio cholerae* (*V. cholerae*) *and V. parahaemolyticus*, which are coastal water associated bacteria (Cui et al., 2015; Shapiro, 2016). Strains of a particular *V. cholerae* lineage, known to cause cholera in humans via drinking of contaminated water, have acquired virulence factors that enable them to cause epidemics (Chun et al., 2009).

In species where recombination is frequent, inferring phylogenetic relationships between strains and lineages can be challenging. This is particularly evident in species where the clonal signal is very weak, and the phylogenetic tree appears as a star-like structure. Such an example, is the genetically heterogeneous *C. concisus* which is found in the saliva and the gastrointestinal tract of humans and is associated with IBD (Kirk et al., 2018). Within highly recombining species like *C. jejuni* there is clonal structure which reflects different ecologies. This is particularly evident in lineages associated with wild birds where isolates from different countries cluster together according to a specific host (Griekspoor et al., 2013; Sheppard et al., 2011). Other bacteria, including *Helicobacter pylori*, have clones that cluster according to geographic location. When *H. pylori* strains pass in a new human host, they evolve at the same time with their host. Studies on *H. pylori* have identified genes that are used as geographical markers to describe the population structure and track human migration movements along the years (Falush et al., 2003).

Successful clones can emerge from a background population and proliferate in a new niche. This can be driven by beneficial mutations that become fixed in the emerging clone and lead to lowering the genetic variation in the population. In contrast, there are species that have very low genetic diversity with only a few single nucleotide polymorphisms (SNPs) differences in their genomes. Typical examples are *Mycobacterium tuberculosis* (*M. tuberculosis*), *Bacillus anthracis* (*B. anthracis*) and *Yersinia pestis* (Achtman, 2008).

Inferring phylogenetic relationships

Early approaches to describe phylogenetic relationships on bacterial species were based on their physical characteristics and metabolic features. The development of various molecular methods has provided insights in the diversity of different bacteria using genes as markers. The most widely used gene marker in phylogenetic analyses, is the small subunit (16S) rRNA gene that was first captured by PCR in 1983 (Muyzer et al., 1993). The first WGS reconstruction was performed in 1995 and since then the number of sequenced genomes has dramatically increased (Fleischmann et al., 1995). Technological advances on sequencing technologies and bioinformatics software have allowed the complete sequence of genomes (Hug et al., 2016). This helped researchers to classify organisms into taxa and species and reconstruct high-resolution phylogenetic trees based on more than a single

gene (e.g. widely used 16S rRNA gene) (Hug et al., 2013). Over the last years, there have been multiple studies on bacterial populations, looking at the population structure and focusing on virulence and/or AMR resistance genes. The concept of core genes (i.e. genes that are shared by all isolates) has made it possible to define and study the population structure of different bacterial strains on all different taxonomic levels: phylum, class, order, family, genus and species. If you extend this to eukaryotes the domain and kingdom levels can also be added. The higher the level in the taxonomic rank, the less core genes are shared between isolates and thus, the less accurate the resolution will be. Defining the group of core genes among different strains is of vital importance to identify conserved genes and their functions and understand their evolution. These genes can be used to study various phenotypic features and infer phylogenetic relationships between strains.

Bacterial host adaptation

Bacteria have adapted to survive in different environmental conditions and different hosts. This adaptive success can be attributed to various structural, metabolic or regulatory mechanisms that bacteria possess (Ryall et al., 2012). Questions remain about how bacterial populations adapt to a new niche following transmission between multiple hosts. This can be the result of a single or multiple transition events between hosts or environments. The genomic adaptations that take place in the bacterial genome that enable them to survive and colonize a new niche are not fully understood and the time scale of this adaptive process is not known. Mutations can slowly alter the genetic variation in the genome and help bacteria adapt into new environments (Thomas and Nielsen, 2005). Additionally, when they are found in a new host, bacteria can speed up adaptation by obtaining large pieces of DNA via HGT in a single event (Vos, 2009). Both strategies have been demonstrated under experimental conditions, where bacteria are typically exposed into different pressures, which are usually changes in the nutrients or exposure to antibiotics (Ryall et al., 2012).

Pathogenic bacteria encounter additional selection pressures compared to non-pathogenic ones, including the host's immune defences and antimicrobials. Studies have found that extraintestinal pathogenic clones of *E. coli* can have higher recombination rates compared to non-pathogenic commensal ones (Rodríguez-Beltrán et al., 2015). The same study shows that recombination rates are extremely

variable in different *E. coli* lineages and are associated with each strain's lifestyle (Rodríguez-Beltrán et al., 2015). In another study, McNally et al. demonstrate a reduction in core genome recombination for the MDR *E. coli* clone ST131 (McNally et al., 2013). Host adaptation in bacteria appears to be multifactorial depending on the host, the environment and the bacterial strain itself.

One way to understand adaptation is to look for genes which confer a competitive advantage in a niche. These genes include virulence or AMR genes, or genes that confer metabolic advantages in colonizing a specific animal host and are usually located on pathogenicity islands. The spread of virulence in *E. coli*, via HGT, likely reflects an adaptive process to a pathogenic lifestyle (Schubert et al., 2009). Similarly, the selection for AMR genes, which are commonly found on a plasmid or other mobile elements, can rapidly spread through the population and transmit between strains and species of the same or different genera (Schubert et al., 2009; Taylor, 1986). The acquisition of AMR genes is driven by the environment, in most cases in agricultural animals and in hospitals, where the antibiotic use is high. Apart from virulence and antibiotic resistance there are other forces that drive bacterial adaptation linked directly with the host. These factors include pH, temperature, exposure to oxygen, the presence and concentration of various nutrients as well as distinct anatomical, histopathological and physiological differences of different animal hosts. In addition, host diet and immune response, in combination with different skin and/or gut bacterial community compositions play a major role in bacterial adaptation.

Host ecology and host associated bacteria

To colonize a new host, bacteria need to modify their genome and metabolic functions in order to adapt and survive. However, questions remain about whether there are specific bacterial lineages that have adapted to specific hosts as a result of long-term evolution over the years, or specific genes that enable adaptation in specific niches (Doolittle and Booth, 2017). To better understand this, we must look in species that have a strong host ecology. The fact that some species have lineages with different lifestyles and are often found in more than one hosts complicates things even further. Understanding the population structure and the gene determinants that are associated with different hosts is one way of looking into it. Core and accessory genome variation can give insights into: (a) genes that are

essential for a group of bacteria (a lineage or a species) and (b) genes found only in strains (or lineages) associated with a particular host. Another way is (c) to measure the effect of recombination in various clones.

A typical example of largely host restricted lineages is demonstrated in *S. aureus* with lineages strictly associated with birds, including poultry, or humans as well as lineages frequently isolated from multiple hosts (Feil et al., 2003). Adaptation to different hosts has been mediated by mobile elements, gene loss, recombination and mutations in the core genome (Lowder et al., 2009; Viana et al., 2010; Murray et al., 2017). Host transitions have been described in *S. aureus* from humans to different agricultural animals. When the human associated poultry lineage CC5 transferred to poultry, it lost genes, associated with human pathogenesis and virulence, and gained novel avian-specific ones (Lowder et al., 2009; Murray et al., 2017). Host adaptations in *S. aureus* have mainly been associated with nutrient availability, with strains from cattle growing better in the presence of lactose compared to human and/or chicken associated strains (Richardson et al., 2018). Some serotypes of the pig-associated bacterium *Streptococcus suis* (*S. suis*) appear to be associated with causing disease in humans, and even though it is likely that these strains have acquired genes associated with virulence to humans, the exact genomic adaptation changes remain unclear (Weinert et al., 2015).

An interesting example of ecology associated population structure is observed in *C. jejuni*. The segregation of lineages with different ecological niches is evident in the existence of livestock, wild bird and environmental associated lineages (Sheppard et al., 2014). For example, some lineages (host specialists) are associated with poultry, cattle or wild birds (Sheppard et al., 2011). In the latter, the segregation of specific lineages to specific wild bird species highlights the strong ecological drivers of evolution in *C. jejuni* which is considered an avian-associated bacterium (Griekspoor et al., 2013). Wild birds pre-date the existence of the modern domesticated chicken (*Gallus gallus domesticus*) and the coevolution between wild birds and *C. jejuni* bacteria appears to be much longer from an evolutionary and adaptive point of view. Experimental infection of wild bird species with different *C. jejuni* lineages, resulted in out-competition of the strains from other sources by the host specialist strains (Atterby et al., 2018). This might reflect strain specific differences in colonization ability which are likely related to physiological factors of

the host (Atterby et al., 2018). However, there are lineages that are occupying multiple niches (host generalists). It might be assumed that when these lineages transmit into a new host, they would be outcompeted by host specialist ones. However, these lineages have been shown to survive in new hosts without being outcompeted by the pre-existing host specialist strains (Atterby et al., 2018). Additionally, there has been enhanced recombination between host generalist and host specialist strains under laboratory conditions (Sheppard et al., 2014) which leads to a likely scenario that, by recombining with host specialists, these lineages acquire the host associated genetic elements that are required to colonize and survive into the new host.

Host adapted lineages and strains are also found in another not thoroughly described species of the genus *Campylobacter*. *C. fetus* consists of three subspecies, *C. fetus* subsp. *fetus* and *C. fetus* subsp. *venerealis* associated with mammals (van Bergen et al., 2005), but also occasionally infecting humans through the food chain production (Wagenaar et al., 2014), and *C. fetus* subsp. *testudinum* associated with reptiles (Fitzgerald et al., 2014). A recent study takes advantage of large sets of whole genome sequenced isolates and suggests that *C. fetus* might have passed from humans to cattle during the first domestication period around 10,500 year ago with adaptive changes observed primarily in the accessory genome (Iraola et al., 2017). More sampling and epidemiological surveillance is required to monitor the prevalence of *C. fetus* in cattle and humans and understand its' transmission patterns and contribution to disease.

Advances in bioinformatics software and pipelines for bacterial genome analyses

The generation of multiple bacterial genomes has opened the possibility for sequence comparisons to study the evolution, within species diversity, outbreaks and virulence/antibiotic resistance. Such comparisons require algorithms and software for reconstructing phylogenies. The Neighbour-joining method is a simple way to reconstruct phylogenetic trees from evolutionary distance data and, even though it is not as accurate, it is reliable enough to give an approximation of the relationships between different strains in less than a minute of computational time (Saitou and Nei, 1987). Following this, one of the most popular ways to reconstruct phylogenetic trees, that is widely used today, is the maximum likelihood (ML)

method, implemented in various software like RAxML (Stamatakis, 2014), PhyML (Guindon et al., 2010) and FastTree (Price et al., 2010). Maximum likelihood is using sequence information to calculate the probabilities for every nucleotide at every position in the alignment that is used as input. However, most bacteria undergo a level of homologous recombination which can vary from species to species (Vos and Didelot, 2009). Reconstructing phylogenies using ML leads to branches on the tree with significant evidence for recombination (Everitt et al., 2014). This skews the true clonal relationships between isolates when trying to reconstruct phylogenies highlighting a need to account for recombination when inferring phylogenetic relationships. ClonalFrameML is a software that accounts for recombination by detecting the recombined regions on each branch of the tree and generates an approximation of the clonal frame. This allows not only the detection of recombining genomic regions, but also the study of the genes in those regions, providing further insights in evolution of several pathogens (Didelot and Wilson, 2015).

Genome wide association studies

Genome-wide association studies (GWAS) have been widely used in human studies, giving valuable insights in the understanding of the genetic basis of phenotypic diversity. GWAS investigate various genetic variants (typically SNPs) within a population of organisms and look for statistically associations matched to a given phenotype (Corvin et al., 2010). If a SNP is identified as significant, then the sequence at the SNP position could be associated with a disease or phenotype trait. The first GWAS in humans was published in 2005 looking for associations between SNPs and macular degeneration, a disease that results in loss of vision (Klein et al., 2005). Many important human diseases have been investigated using GWAS (Welter et al., 2014). Although, GWAS proved to be a vital tool to use in microbiology, it was not used in bacteria until the start of large-scale genome sequencing (Read and Massey, 2014).

The first GWAS study in bacteria was published in 2013 looking for associations in genetic factors associated with host adaptation in *C. jejuni* (Sheppard et al., 2013). Since then, more studies using GWAS in bacteria were published highlighting the power of this method in characterizing phenotypes at a population level (Farhat et al., 2013; Laabei et al., 2014; Alam et al., 2014). The last 10 years more updated and modified GWAS methods have been introduced to implement this approach into

bacterial genetics (Chen and Shapiro, 2015). Some examples of software pipelines and tools for performing GWAS are: SEER (Lees et al., 2016), treeWAS (Collins and Didelot, 2018), dbGWAS (Jaillard et al., 2018), bugWAS (Earle et al., 2016) and bacGWASim (Saber and Shapiro, 2020). Several studies have applied GWAS methods to identify genetic variants associated with virulence, antibiotic resistance and host-associated factors in bacterial populations of closely related strains (Pascoe et al., 2015; Alam et al., 2014; Berthenet et al., 2018). However, many challenges exist when applying GWAS to bacteria. When designing and performing GWAS it is important to consider the population structure of the species used for analysis. Some bacterial species undergo high levels of recombination and that creates a problem to structure genetic variation. For example, it is difficult to differentiate the associations that result as part of the clonal frame from the adaptive ones and identify those SNPs that are associated with a phenotype and cannot be explained by the phylogeny. Further validation of genotype-phenotype relationship could be achieved with high-throughput phenotypic assays, depending on the functional mechanism of the genes or the genomic regions identified by GWAS, with *in vitro* and *in vivo* assays. Bacterial GWAS is a very recent introduction into the field and although it provides a powerful tool for phenotypic characterization at a population level, the above challenges need to be addressed (Read and Massey, 2014).

Dating molecular phylogenies

Dating the ancestry of a lineage is important to understand evolution. Researchers can study the emergence of lineages, linking them to known biological, geological and/or historical events. Knowing the ancestry of different lineages provides information about the speed of diversification of different groups and the rates of the molecular clock (Kuo and Ochman, 2009). In palaeontology, dating a lineage can be done directly when identifying a fossil and calculate divergence time on a phylogeny by comparing morphological and molecular estimates (Smith et al., 2006). This cannot be applied in bacteria where no fossil record is available. Instead, molecular clocks have been used to estimate divergence times, in bacteria as well as in viruses (Sheppard et al., 2010; Falush et al., 2001; Bromham and Penny, 2003). To calibrate molecular clocks in bacteria, one could calculate the number of mutations that accumulate over evolutionary timescales, assuming those occur at a constant rate. However, the mutation rates for natural populations of bacteria are

usually not reliable. Furthermore, in highly recombinogenic bacteria, recombination can shuffle sequence polymorphisms in large parts of the genome resulting in phylogenies with false branch lengths, inaccurate dating to the most recent common ancestor (MRCA), large confidence intervals, loss of the molecular clock estimate and overestimation of the substitution rate heterogeneity (Schierup and Hein, 2000). However, using the isolation date of each isolate it is possible to attempt to date divergence times of nodes across a phylogeny. This has potential to provide information about the emergence of new lineages and/or their associated novel genetic traits (Rieux and Balloux, 2016). This can also be applied to non-recombinogenic bacteria like *M. tuberculosis* (Eldholm et al., 2015). Obviously, in order to accurately perform molecular dating, isolation dates are needed. Specifically, information about the different time points of sampling, either months or years, for every isolate. This will allow the study of molecular phylogenies on a natural timescale assuming there are significant changes in the genomes between the sampling times, and that contemporary isolates share ancestry within the sample period. TempEst, a tool that compares the genetic divergence that we see on a phylogenetic tree through sampling time, has been recently introduced (Rambaut et al., 2016). By using a regression approach, this tool allows researchers to interrogate their data for temporal signal before applying more sophisticated dating methods. These methods include Bayesian phylogenetic inference, implemented in the well-known Bayesian Evolutionary Analysis by Sampling Trees (BEAST) software which applies advanced molecular clock rate models and Markov chain Monte Carlo (MCMC) algorithms to calculate divergence time, coalescence and phylogeography (Drummond et al., 2012). The updated version BEAST2.5 offers more packages and post processing tools including hierarchical multispecies coalescent models for tree estimation and more epidemiological models to investigate rapidly evolving infectious disease pathogens (Bouckaert et al., 2019). BEAST requires computer power, with large-scale sequence analyses requiring long computational time, so other software offer alternative ways of dating molecular phylogenies. Least-square criteria and algorithms provide fast molecular dating of phylogenies with results being highly comparable with those from BEAST in some microbial datasets (To et al., 2016).

Many pathogenic species can colonize multiple hosts, with emerging clones following transitions between different hosts. Estimating time divergence between

these isolates would provide information about the timeline of different host switches and possibly link to events associated with those switches. The fore mentioned software has been particularly useful in understanding pathogen dynamics, emergence and spread of pathogenic bacteria in human and animal populations. This has been extensively tested in viruses with broad sampling metadata like HIV (Faria et al., 2014), Influenza (Worobey et al., 2014) and Ebola (Gire et al., 2014) to monitor their molecular evolution over time. In recombinogenic bacteria, the temporal signal can be distorted by introduction of additional divergence between two isolates. However, removing highly recombinant regions from the genome is a way of handling that problem. Molecular dating has been used to estimate host switching events between humans and bovids in *S. aureus* (Weinert et al., 2012) but also cross-species transmission events between different hosts (Richardson et al., 2018). The emergence of a *S. suis* zoonotic clade has been estimated in the beginning of the 20th century, coinciding with the intensification of animal agricultural systems in the pig industry (Weinert et al., 2015). A study in *C. fetus* has found that strains of that species jumped from humans to cattle following the domestication of cattle 10,500 years ago (Iraola et al., 2017). An attempt to attribute the source of infection of *C. jejuni* lineages, that occupy multiple niches, could not pinpoint the source to a specific host, as the host signal is blurred, due to strains undergoing multiple host jumps (Dearlove et al., 2016). The recently developed software, BactDating, was introduced to deal with molecular dating in bacteria using Bayesian estimation of the molecular clock rate (Didelot et al., 2018). This method includes test functions that are compatible with outputs of software that detect recombination in bacteria like ClonalFrameML (Didelot and Wilson, 2015) and Gubbins (Croucher et al., 2015).

Multihost pathogen emergence

The majority of human or animal pathogens can infect more than one host species. Some of those multihost pathogens can be transmitted by multiple hosts, which makes them ecological generalists and those pose a high risk in public health (Wardeh et al., 2015). Zoonoses, i.e., infections transmitting from animals to humans, account for >60% of all human pathogenic diseases, causing not only major public concern but also a massive burden on global economies (Woolhouse, 2001). The causative agents behind zoonoses commonly include viruses, bacteria, parasites and fungi. The currently on-going Coronavirus pandemic, caused by

COVID-19 virus, is an example of how terrible the effect of a sudden pandemic can be not only on public health, but also on the global economy, with predictions giving shocking estimates about the upcoming economic losses (Dong et al., 2020). COVID-19 has to date infected >16 million people causing more than half a million deaths worldwide, with numbers still increasing in the US and Latin America (Dong et al., 2020). The epidemics of HIV/AIDS emerged following a chimpanzee to human transmission and have killed millions of people and even though various of treatments are available, the estimated cost is ~$2,000 per patient per year (Sharp and Hahn, 2011; King et al., 2006). In 2003, the first novel human coronavirus took place, with bats being the most likely reservoir of the virus (Bolles et al., 2011). Also known as severe acute respiratory syndrome (SARS), this epidemic killed less than 1,000 people (being less contagious than SARS-CoV-2), but its' effect on the economy was devastating costing the global economy around $40 billion (King et al., 2006). The influenza virus is notorious for causing outbreaks/epidemics/pandemics over the last centuries, perhaps even earlier in history (Girard et al., 2010; Nelson and Holmes, 2007).

Bacterial epidemics have been a major problem for the public, with more and more epidemics being characterized, owing to the new tools of the genomic era (Raoult, 2014). This is an alarming problem as infections, associated with resistance to antibiotic treatments, are worryingly increasing. Hospitals, medical centres and other health care environments, where patients visit frequently, have been associated with outbreaks. Epidemic waves of hospital-associated methicillin-resistant *S. aureus* have been very common over the last 20 years (Li et al., 2012). Numerous of anthrax-related epidemics and outbreaks in humans and animals have been described since 1950 with *B. anthracis* as the causative agent (Bales et al., 2002). Food-borne bacterial pathogens have been widely known to cause outbreaks. Enterohaemorrhagic serotype *E. coli* O157:H7 was first reported in China in 1990 and has gained a lot of publicity since the outbreaks in US and Japan in 1993 and 1996, respectively (Shao et al., 2011). Most food-borne diseases are zoonoses and have *Campylobacter* and *Salmonella* as their dominant causing agents worldwide (Thorns, 2000). Other epidemics and outbreaks have been described, like the ones caused by group A *Streptococcus* (Beres et al., 2010), *S. enterica* subsp. *enterica*, *Salmonella enteritidis* and *Salmonella typhimurium* (Bäumler et al., 2000; Bush et al., 2011). Ongoing anthropogenic changes to the

planet, like deforestation and the destruction of wildlife habitats along with the increase in meat, eggs and dairy products on the market, have brought wildlife, livestock and humans closer together increasing the emergence of pathogens which are causing epidemics/pandemics.

Antimicrobial resistance and the dangers ahead

Since the discovery of AMR for the first time in 1924 against a drug known as salvarsan, that was widely used to treat syphilis between 1910 and 1914 , many serious bacterial-associated outbreaks, have been reported (Bush et al., 2011). When the bacteria causing an outbreak are susceptible to antibiotics, patients can be easily treated, for example in the *E. coli* outbreak in Germany in 2011 (Frank et al., 2011). However, when the underlying pathogen is associated with high antibiotic resistance, then the outbreak/epidemic can be very problematic. In recent decades many bacterial pathogens have acquired resistance to different classes of antibiotics. This has been affected by the extended use of antibiotics in human and veterinary medicine but also in modern agriculture with the use of antibiotics as growth promoters in animal feed (Castanon, 2007). Antibiotic resistance can be a large problem in hospital settings where patient-to-patient transmission is common and can lead to rapid spread of MDR bacteria (Bush et al., 2011). Pathogenic bacteria can acquire AMR genes from the environment through HGT, with many of those genes located on transposons, plasmids and other mobile elements (Boerlin and Reid-Smith, 2008). The gut can act as a gene pool for AMR transmission events between commensal and pathogenic bacteria (Salyers et al., 2004).

Numerous AMR genes that confer resistance to antibiotics have quickly spread among bacterial species. The new Delhi metallo-β-lactamase-resistance gene (NDM-1) conferring resistance to broad spectrum β-lactams and carbapenems, has spread worldwide due to frequent international travel and its' high transmissibility potential (Giske et al., 2012; Datta et al., 2017). Usually located in plasmids, *bla*$_{\text{NDM-1}}$ has been detected in various human pathogenic bacteria including *E. coli*, *Klebsiella pneumoniae* and *Acinetobacter baumannii* (Poirel et al., 2011). Another example of dissemination of AMR genes across different bacterial species is the *tet*-related genes, conferring resistance to tetracycline. Plasmid-mediated transfer of *tetO* gene from a Gram-positive to Gram-negative bacteria has been described (Taylor et al., 1983; Batchelor, 2004). This gene is highly prevalent in *Campylobacter* species, circulating among isolates from animals, humans and

urban effluents (Mourkas et al., 2019). In addition, other tetracycline resistance genes including *tet*(X) have been found in MDR bacterial strains of various pathogenic bacteria including *E. coli, K. pneumoniae* and *Enterobacter* (Leski et al., 2013). Infections caused by *K. pneumoniae* are a common problem in intensive care units (ICUs). Although resistance against carbapenems was known for 22 years, it was not until recently that it was recognized as a major health problem. Carbapenem resistance, caused by the gene KPC (*Klebsiella pneumoniae* carbapenemase), encoding for an enzyme that can break down carbapenems, has been also located on a plasmid which makes it easier for dissemination from one bacterium to another (McKenna, 2013).

Increasing AMR has become one of the world's most serious contemporary public health concerns. The World Health Organization (WHO) has placed AMR as one of the most serious threats in the years to come as more and more bacteria are becoming resistant to all known antibiotics. The last generation penicillin drugs, the carbapenems, are already ineffective in many countries like Greece, Spain and India. It is quite alarming that there have been no new drugs that target the Carbapenem-resistant *Enterobacteriaceae* bacteria for the past 40 years. It has become apparent that modern medicine is under crisis and that without the discovery of new drugs or combination therapies, AMR will become one of the biggest global challenges to public health in the near future (Coates, 2019).

Aims of this thesis

The work in this thesis has attempted to advance our understanding of bacterial adaptation in different hosts and environments by investigating genomic changes and linking them, when applicable, to phenotypes. This work has used the power of comparative genomics to examine: (i) differences in the core and accessory genome and (ii) the effect of recombination in the evolution of *Campylobacter*.

1. In chapter 2, bacterial genomes from various *Campylobacter* species were analysed to answer questions related to what maintains species as discrete entities and the extent of genetic exchange between them. Subsequently the following questions will be investigated:
    - Do *Campylobacter* share more accessory and/or core genetic elements when found in the same host?
    - Can distant *Campylobacter* species recombine, or are there other barriers to recombination?
    - Are there any hotspots of recombination and what is the proportion of the genome that recombines?
2. Chapter 3 focuses on the impact of intensive livestock production on the evolution of host-associated *C. jejuni* lineages. In this chapter, comparative genomics, linked to phenotype studies, were used to investigate the genomic adaptations and timescale of the emergence of cattle-associated lineages. The following questions will be investigated:
    - How do cattle-associated *C. jejuni* lineages emerge?
    - What is the timescale of this emergence?
    - What are the genomic elements associated with adaptation into cattle?
    - Can genotype-phenotype be linked mechanistically?
3. In chapter 4, the genomic variation associated with AMR was investigated in *Campylobacter* isolates from humans, livestock and sewage to answer questions related to the spread of AMR among source and sink populations. The following questions will be investigated:
    - What is the MDR profile of *C. jejuni* and *C. coli* isolates?
    - How is AMR distributed across *Campylobacter* source/sink populations?
    - Are there species or lineages that harbour more AMR?
    - Is there movement of AMR between animal, human and environmental gene pools?

4. Finally, chapter 5 presents the use of genomics as a tool to identify lineage-specific genes to monitor bacterial load in animal infection experiments. The following questions will be investigated:
    - Do distinct *C. jejuni* lineages have their own pool of accessory genes?
    - Can genomics be utilized to identify lineage-specific gene markers that can later be used for monitoring bacterial load in *in vivo* animal experiments?

Chapter 2

Interspecies recombination in agricultural *Campylobacter* is influenced by the song (gene) and the singer (strain)

Commentary text

The work in this chapter provides a detailed analysis of genome-wide genetic diversity within the genus *Campylobacter*. Thirty *Campylobacter* species from a range of hosts and environments, were analysed in order to answer questions related to evolutionary forces that maintain what is known as species. Quantifying intra- and interspecies recombination showed that species share more genetic elements when they are found in the same host or environment. Genomic analyses identified recombination hotspots, that were mapped to genes with various functions. These genes represent potential adaptations of species that cohabit the same niche. The work in this chapter presents a novel way of looking at the extent of genetic exchange between closely related *Campylobacter* species and explores the limits of bacterial species boundaries. The statement of authorship for this chapter can be found in the Appendix, supplementary form SF1.

Abstract

Evolutionary analyses of bacteria often consider lineages and species that inhabit different niches. For example, within the genus *Campylobacter* there are species that inhabit the gut of different animals and species that inhabit different niches within a single animal. The maintenance of these species as discrete entities depends on barriers to genetic exchange between them. These can be physical - with species inhabiting different niches, or adaptive - implying selection against hybrid lineages, but quantifying the relative importance of these barriers can be challenging. Considering genes, rather than lineages, as units of selection provides a theoretical solution to this. While understanding clonal population structure and phylogenetics remains important, new theoretical approaches consider the genes that underlie the collective functions of a microbiome (songs) rather than the lineages in which they are found (singers). Here, >600 genomes of multiple *Campylobacter* species isolated from birds, mammals and reptiles were analysed. By characterizing interspecies core and accessory genome recombination in isolates from the same and different hosts, this chapter quantifies the extent to which genes, rather than lineages, inhabit the niche. Specifically, for some species pairs there was ~0.6 times more recombination between cohabiting isolates than host segregated ones. By broadly defining the limits of interspecies recombination and the function of mobile genes, work in this chapter provides real-world data by which to interrogate influential theories about the levels of the biological hierarchy (genes, lineages, species) at which selection operates to maintain what is known as 'species'.

Introduction

Industrialisation of farming practices has dramatically changed the distribution of plants and animals on Earth. These radical anthropogenic ecological changes, niche expansion and subsequent influx of bacterial species has removed some of the physical barriers to recombination. Bacterial species boundaries can be blurred by recombination between distantly related bacterial species (Shapiro et al., 2016; Doolittle and Zhaxybayeva, 2009), challenging the traditional Linnaean taxonomy and biological species concept used to describe most Eukaryotic species (Dobzhansky, 1935; Mayr, 1942). The dramatic niche expansion greatly increases the effective population size (Ne) of bacterial populations, which leads to increased genetic variation. Genetic variation is maintained through neutral selection (genetic

drift) or ecology-driven selection (Sheppard et al., 2018), giving identification of fuzzy species that do not form clear and distinct clusters (Dobzhansky, 1935; Mayr, 1942; Hanage et al., 2005; Hanage, 2013; Mallet et al., 2016).

However, phylogenetic clustering based on core genes (shared by all isolates in a comparison) tend to support a clonal population structure (species), which implies that there are still barriers to recombination. For recombination to occur there must be homology between a recipient genome and the donor DNA. Restriction-modification systems act as defence mechanisms against the uptake of foreign DNA (mechanistic barriers) (Thomas and Nielsen, 2005; Eggleston et al., 1997). Bacteria must also inhabit the same space for the exchange of genetic material (ecological barriers). Novel genotypes can also be outcompeted by fitter predecessors (adaptive barriers) purging the population of sub-optimal recombination events (Zhu et al., 2001). Through modelling genetic admixture (Lawson et al., 2012) recombination can be quantified and hot spots of variation associated with phenotype differences, identified (Sheppard et al., 2013; Yahara et al., 2017; Berthenet et al., 2018; Yahara et al., 2016; Yahara et al., 2018; Yahara et al., 2014).

Host-associated bacteria must strike a balance with the pre-existing microbiota (Davenport et al., 2017) to help evade the host immune response (Tuchscherr et al., 2011) and utilise the available nutrients (Sheppard et al., 2013). A *plug-and-play* approach to bacterial accessory genomes (Young, 2016; McInerney et al., 2020), where diversity in bacterial phenotypes can be provided by a mobile pool of genes that are readily gained and lost, enables the acquisition of rapid adaptive genomic changes that can be spread through the population via recombination (Mourkas et al., 2020). This results in different strains carrying unique selections of accessory genes drawn from the species' pan-genome (Vos and Eyre-Walker, 2017). These selfish elements reproduce and spread within and between lineages (Werren, 2011), sometimes in conflict with what is best for the individual (Doolittle and Papke, 2006). Short generation times allow selection to act quickly on bacterial populations, encouraging rapid adaptation and expansion into new niches (Holt, 2009).

In agriculture, a reduction in physical recombination barriers has led to an increase in the emergence of zoonotic diseases, where host-restricted species can broaden their infectious host profile and encounter other bacterial species. Cross-species

introgression has shaped the population structure of *Campylobacter*, a common gastroenteritis pathogen. *C. jejuni* lineages are often associated with a single host (specialists) (Griekspoor et al., 2013; Mourkas et al., 2020; Sheppard et al., 2011), or multiple hosts (generalists) (Dearlove et al., 2016). A single host can simultaneously carry multiple lineages (Colles et al., 2008) possibly occupying different sub-niches within that host (Colles et al., 2015). Quantification of interspecies recombination between *C. jejuni* and *C. coli* isolated from agricultural animals suggests that up to 23% of the *C. coli* genome has been acquired from *C. jejuni* (Sheppard et al., 2013).

In this chapter, over 600 *Campylobacter spp.* isolates (from 30 designated *Campylobacter* species) were analysed. Isolates were selected to represent genus-wide genetic diversity from a range of host species, which provide the opportunity to consider the relative influence of two evolutionary pressures: (1) Do *Campylobacter* share more accessory and/or core genetic elements when found in the same host? (2) Can distant *Campylobacter* species recombine, or are there other barriers to recombination? *Is it the song (genes) or the singer (bacterial species)?*

Results

Host ecology influences the population structure of the *Campylobacter* genus

In this chapter, genomes from publicly available databases were used, representing genus-wide diversity within *Campylobacter*. These were augmented with environmental isolates from the closely related *Arcobacter* and *Sulfurospirillum* species to evaluate diversity within the *Campylobacteraceae* family (Figure 2.1). In total, 631 isolates from 30 different *Campylobacter* species (Figure 2.2) and 64 different sources, isolated from 31 different countries between 1953 and 2016 were included (Appendix, Table S2.1). Among the 631 isolates, 361 were *C. jejuni* and *C. coli* and could be classified according to 31 Sequence Type complexes (ST-complexes) based upon sharing four or more alleles at seven housekeeping genes defined by multi-locus system typing (MLST) (Appendix, Table S2.1) (Dingle et al., 2001) and were representative of the known diversity in both species (Mourkas et al., 2020; Sheppard et al., 2011).

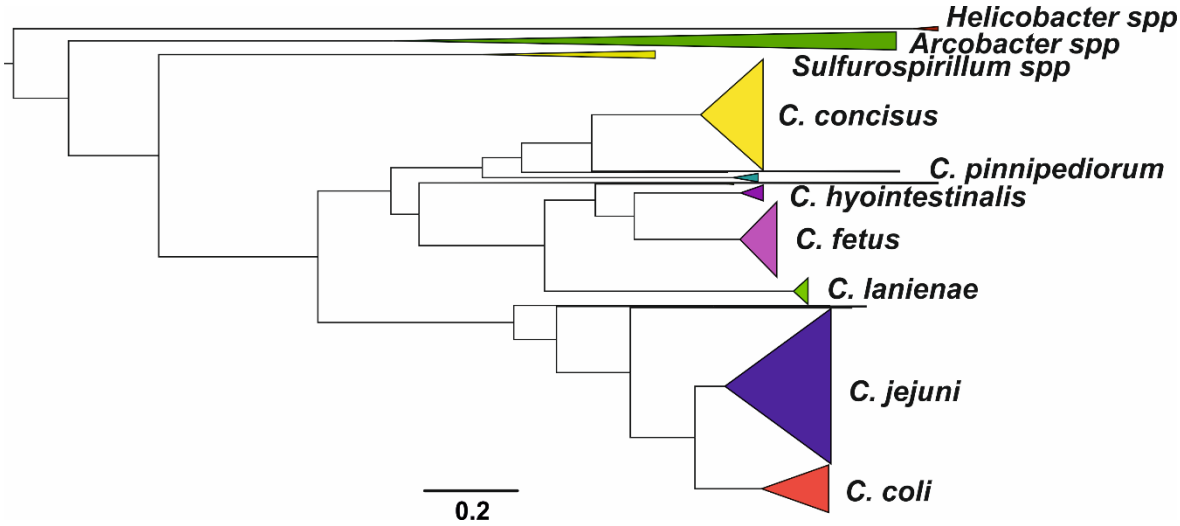


Figure 2.1. Population structure of the *Campylobacteraceae* family. Phylogenetic tree of 662 isolates that belong to the *Campylobacteraceae* family with *Helicobacter pylori* used as an outgroup. The tree was reconstructed using a gene-by-gene concatenated alignment of 603 core genes shared by >95% by all isolates and an approximation of the maximum-likelihood algorithm (ML) implemented in RAxML, with the major species indicated next to the associated sequence cluster. The scale bar indicates the estimated number of substitutions per site.

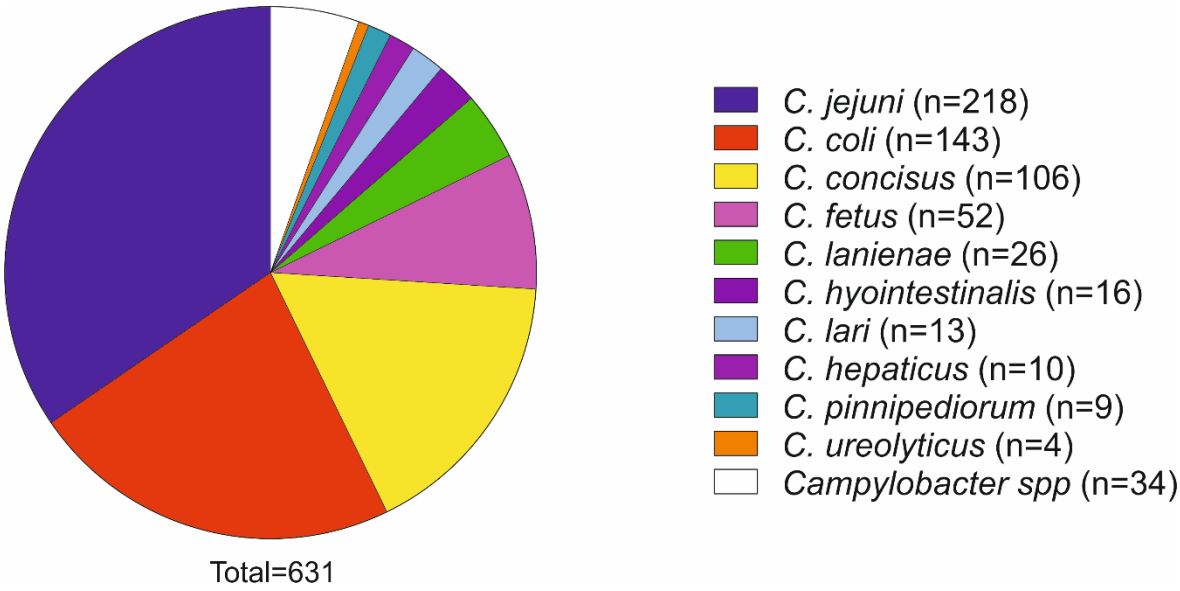


Figure 2.2. Species attribution of 631 genomes of the genus *Campylobacter*. Different colours correspond to main species with number of isolates greater than three.

The obligate human pathogenic *C. concisus* species (n=106 isolates), consisted of the two phenotypically indistinguishable but genetically distinct clusters, genomospecies I (GSI; n=32) and genomospecies II (GSII; n=74), as previously described (Kirk et al., 2018) (Appendix, Table S2.1). The collection also included more than 50 *C. fetus* (n=52) isolate genomes, consisting of *C. fetus* subsp. *fetus* (n=8), *C. fetus* subsp. *venerealis* (n=23) and *C. fetus* subsp. *testudinum* (n=21) (Appendix, Table S2.1) (Iraola et al., 2017).

A maximum-likelihood phylogeny of the *Campylobacter* genus was reconstructed on a gene-by-gene concatenated sequence alignment of 820 gene families shared by >95% of all isolates, with a core genome of 903,753 base pairs (Figure 2.3). The phylogeny illustrates the broad source diversity of isolates from *C. jejuni* and *C. coli* species (ecological generalists) (Figure 2.3). Single-species trees showed a clear population structure with isolates clustering into 28 ST-complexes in *C. jejuni* and three clades in *C. coli* (Figure 2.4). A distinct separation was observed for both mammalian *C. fetus* subsp *fetus* and *C. fetus* subsp *venerealis* and reptile *C. fetus* subsp *testudinum*, as well as for the genetically heterogeneous *C. concisus* GSI and II, for *C. fetus* and *C. concisus*, respectively (Figure 2.4) as previously described (Iraola et al., 2017; Kirk et al., 2018).

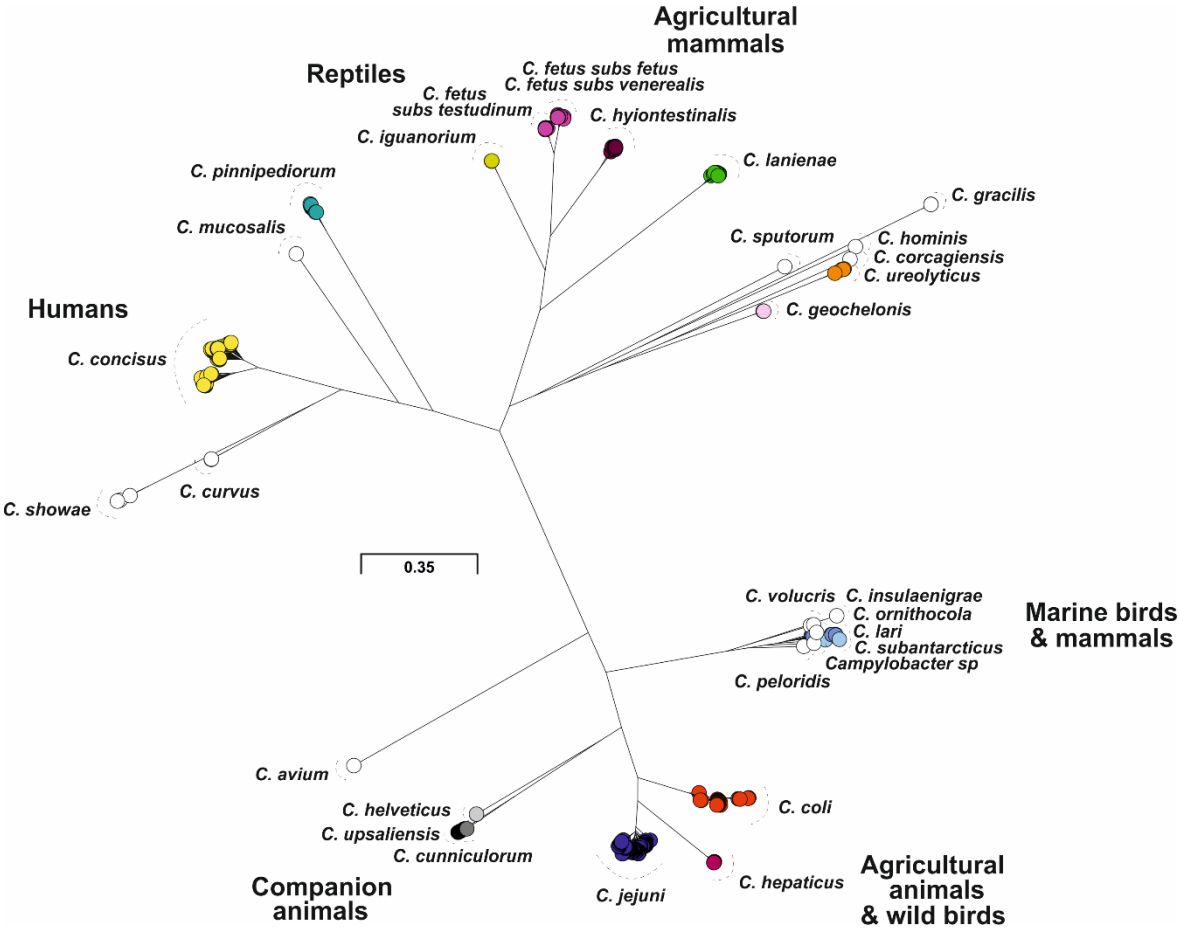


Figure 2.3. Population structure of the *Campylobacter* genus. Phylogenetic tree of 631 *Campylobacter* isolates from all available species. The tree was reconstructed using a gene-by-gene concatenated alignment of 820 core genes shared by >95% by the isolates and an approximation of the maximum-likelihood algorithm (ML) implemented in RAxML, with the species and the most common associated sources indicated next to the associated sequence cluster. The scale bar indicates the estimated number of substitutions per site.

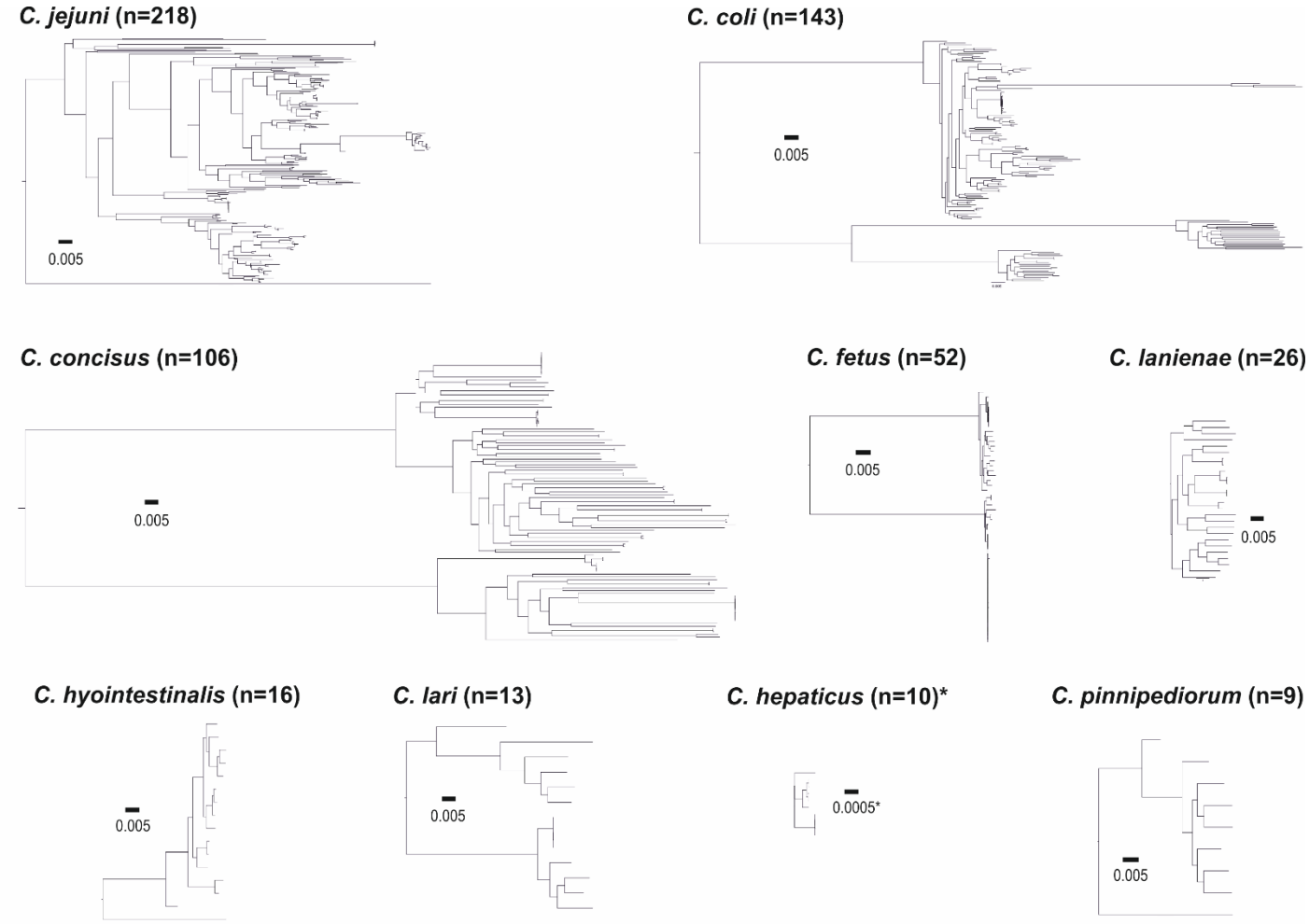


Figure 2.4. Core genome species trees. Single-species trees for nine *Campylobacter* species with >4 isolates demonstrating the diversity among species. The scale bars indicate the estimated number of substitutions per site. (*) The scale for the tree corresponding to *C. hepaticus* is 10 times smaller than the rest.

The *Campylobacter* genus tree also includes uncommon species which appear to be restricted to one host or environment, including *C. iguanorium* (Gilbert et al., 2015) and *C. geochelonis* (Piccirillo et al., 2016) (reptiles), *C. lanienae* (Logan et al., 2000) (pigs), *C. hepaticus* (Van et al., 2016) (chicken liver), *C. lari* group (Miller et al., 2014) (marine birds and environment) and *C. pinnipediorum* (Gilbert et al., 2017) (seals) species, most of which discovered recently (Figure 2.3, Figure 2.5). Single-species trees showed lower diversity for species with n>9 probably due to lack of sampling, with *C. hepaticus* having the lowest diversity (Figure 2.4) with 8/10 genomes associated with isolates from the same outbreak (Van et al., 2016). Two clades were observed in *C. lari* (Figure 2.4) which could correspond to previously described subspecies based on 16S rRNA sequencing (Debruyne et al., 2009).

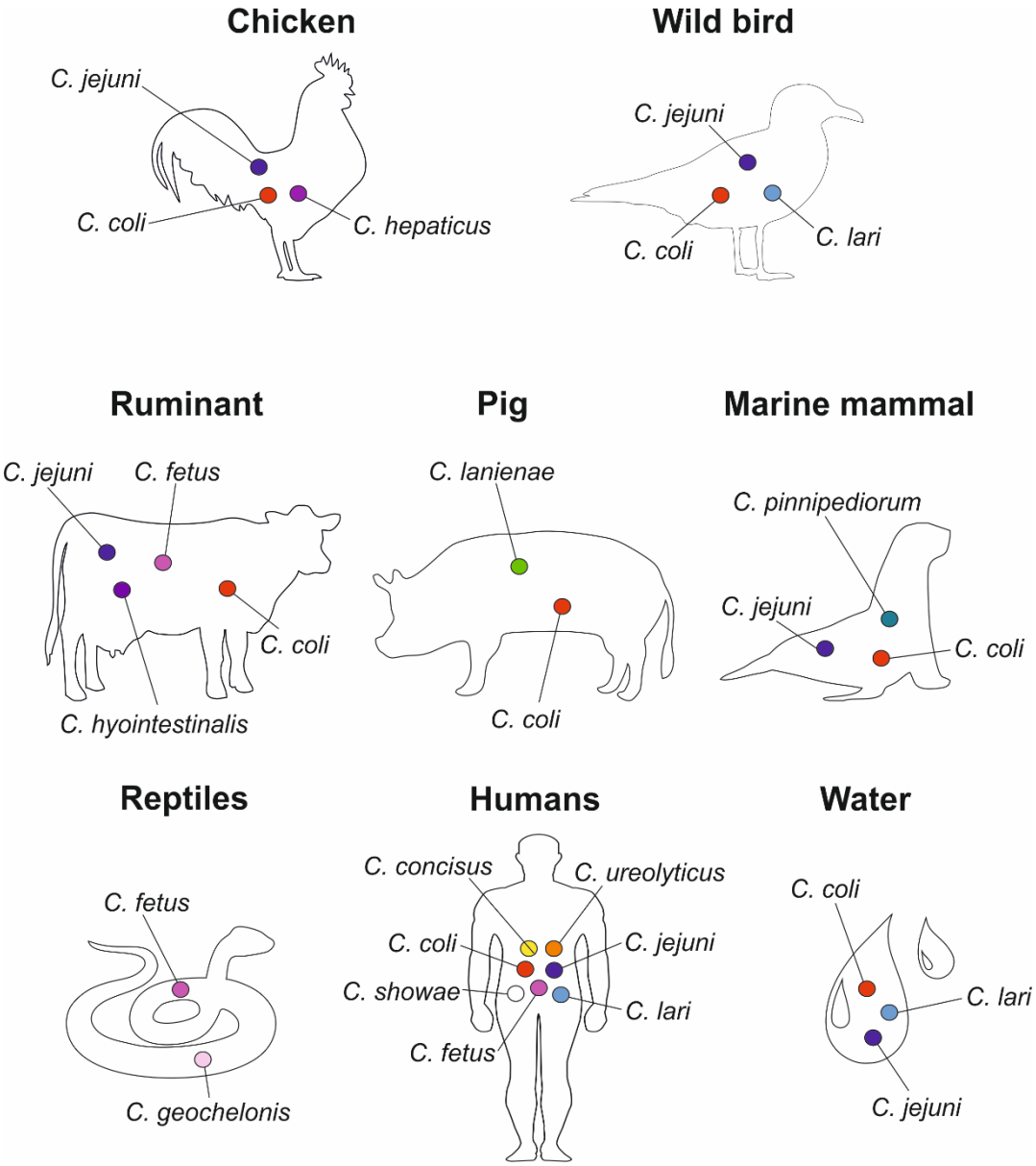


Figure 2.5. Overview of host-associations of *Campylobacter* species. The most common sources where different *Campylobacter* species (n≥3) have been sampled. Each *Campylobacter* species is represented as a central node, with different colours corresponding to different species.

More than 1/3 of the isolates (38.35%; 242/631), were from nine different *Campylobacter* species and isolated from an agricultural animal (Appendix, Table S2.1). The presence of both host generalist and specialist *Campylobacter* lineages and species emphasizes ecology as an important factor associated with the evolution of the *Campylobacter* genus. Clinical isolates were predominantly *C. jejuni* (27.52%; n=60/218) and *C. coli* (14.68%; n=32/218) or *C. concisus* (44.5%; n=97/218; Appendix, Table S2.1). However, at least one isolate from nearly half of all *Campylobacter* species (44.83%, n=13/29) was from a human clinical sample, hinting at their potential to infect human populations (Figure 2.5, Appendix, Table S2.1).

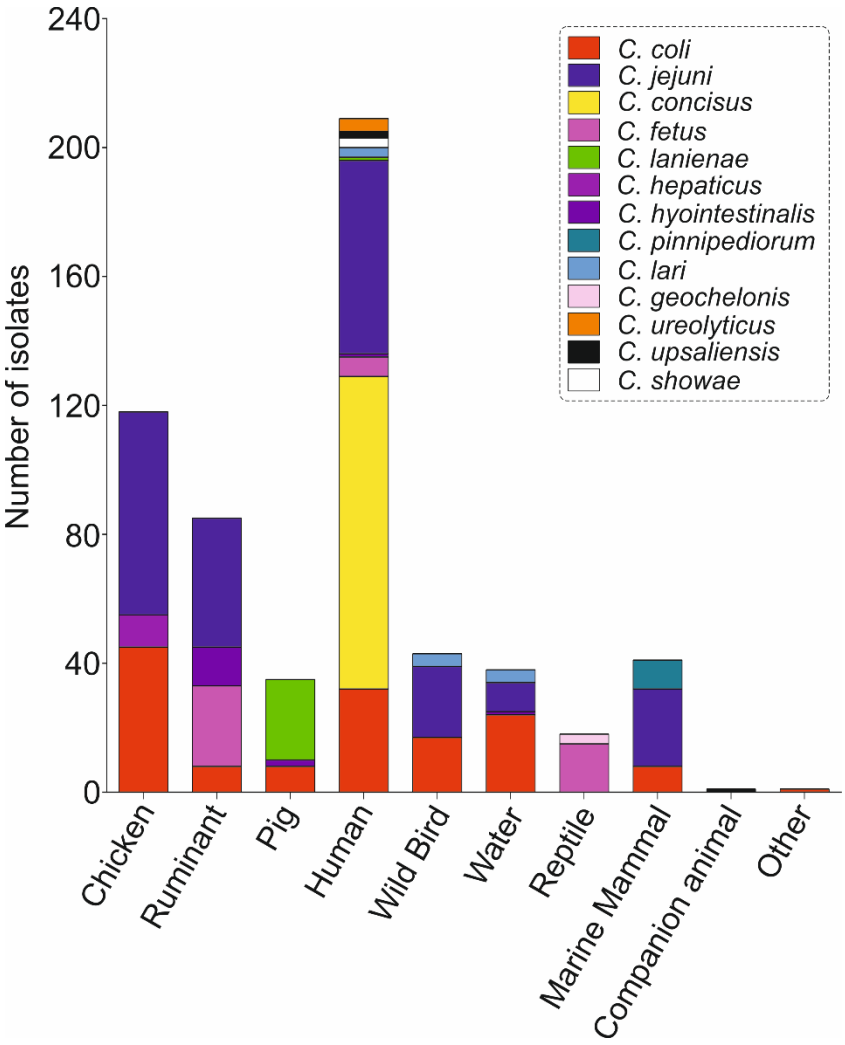


Figure 2.6. Abundance and diversity of *Campylobacter* species in each host and environment. The number of isolates is shown on the *y* axis. Different colours correspond to main species with n≥3.

Pangenome genome content varies between species in the genus *Campylobacter*

Inter- and intraspecies-specific genome content variability was investigated by comparing the whole genome sequences of all 631 genomes (Figure 2.7). Genome size varied between 1.40 and 2.51 Mb (mean 1.73) and the number of genes (per isolate) ranged from 1,293 to 2,170 (mean 1,675) (Figure 2.8 top).

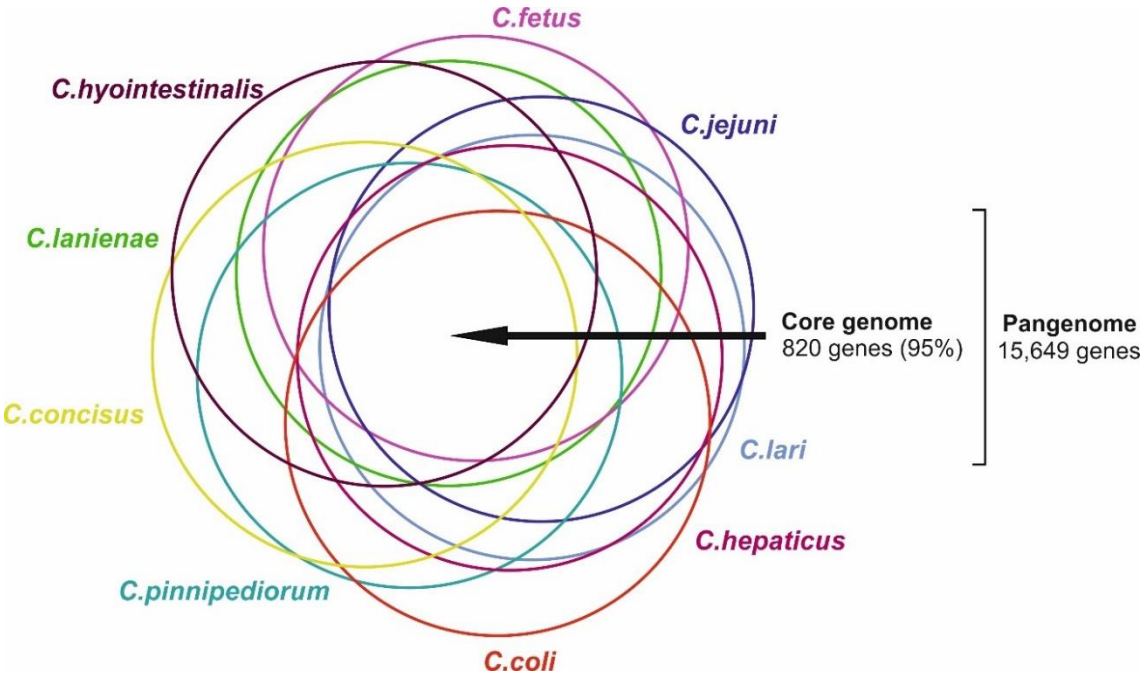


Figure 2.7. Schematic representation of pangenomes between different *Campylobacter* species (where n>4 isolates). The number of core genes shared by all species is illustrated in the middle of the venn diagram.

After removal of allelic variants, the pangenome consisted of 15,649 unique genes, found in at least one of the 631 isolates (Figure 2.7). A number of 820 genes (5.24 % of the pangenome) was shared by >95% of all isolates (core genome), across 15 species. (Figure 2.7). Species with a small number of isolates (n<3) were excluded from gene quantitative analysis. Work in this chapter identified variation in the complement of core genes ranging from 1,116 in *C. lari* to 1,700 in *C. geochelonis*, for each of the 15 species (Figure 2.8 middle). Large differences were noted in the size of accessory genomes (14,829 genes representing 94.76% of the pangenome), with *C. concisus* (mean: 981 genes), *C. hyointestinalis* (mean: 946 genes), *C. showae* (mean: 1,160 genes), *C. geochelonis* (mean: 1,021 genes) and *C. fetus* (mean: 912 genes) containing the highest average number of accessory genes (Figure 2.8 bottom).

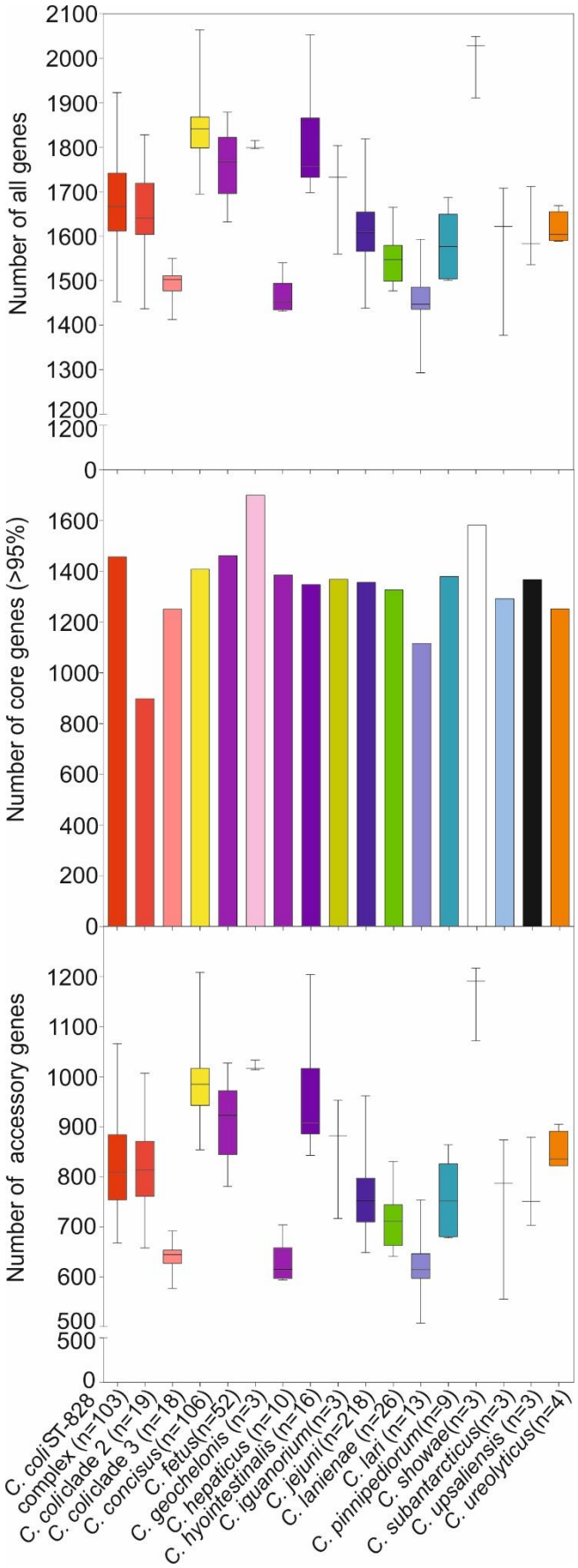


Figure 2.8. Pangenome analysis of the *Campylobacter* genus. Overall distribution of the total number of genes per isolate for each *Campylobacter* species (where n≥3 isolates). The total number of genes (top), core genes (middle) and accessory genes (bottom) is illustrated in all three plots. The number of genes (for total and accessory) is shown as boxplots (min to max).

Campylobacter share more genetic elements when found in the same niche

Exploring the host colonization range in the genus showed that some hosts are colonized by multiple lineages of the same or different *Campylobacter* species (x_1, y; Figure 2.9). Evidence of genetic exchange between *Campylobacter* lineages of the same (Mourkas et al., 2020) or different species (Sheppard et al., 2013) has been described before. The chapter's hypothesis was that there is more genetic similarity between strains occupying the same niche ($x_1 \rightarrow y > x_2 \rightarrow y$; Figure 2.9). To test this, work in this chapter quantified: (1) accessory genome variation, (2) core genome allelic variation and (3) the effect of recombination, in a set of 297 *Campylobacter* strains from 15 different species, isolated from chickens, cattle, pigs, wild birds and environmental waters (Appendix, Table S2.1).

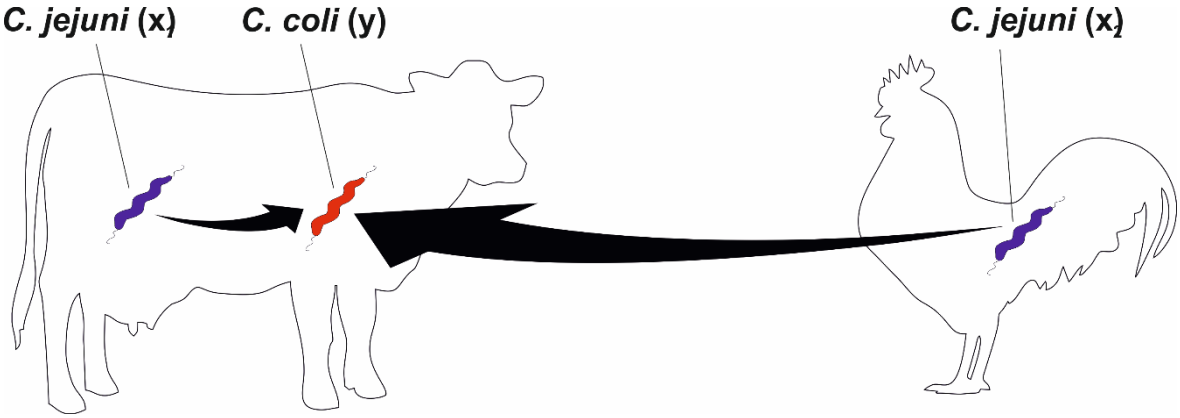


Figure 2.9. Recombination when species co-habit hypothesis. Figure represents the chapter's hypothesis and shows the relationships between *Campylobacter* species, *C. jejuni* (x_1, x_2) and *C. coli* (y), when found in the same or in different hosts.

(1) Agricultural-associated *Campylobacter* species harbour more AMR genes

The Pangenome Neighbour Identification for Bacterial Populations (PANINI) software (Abudahab et al., 2019) was used to investigate shared accessory gene content between distant *Campylobacter* species. Two-dimensional clustering was consistent with the population structure of the tree, reflecting similarities on core and accessory gene content (Figure 2.10). Functional annotation of 14,829 accessory genes showed that 71% (10,561) were encoding hypothetical proteins of unknown function due to the lack of homology with well-characterized genes (Figure 2.11) (Pascoe et al., 2019). The rest of the genes were related to metabolism, DNA

modification, transporters, virulence, inner membrane/periplasmic, adhesion, regulators, metal transport and antimicrobial resistance (Figure 2.11).

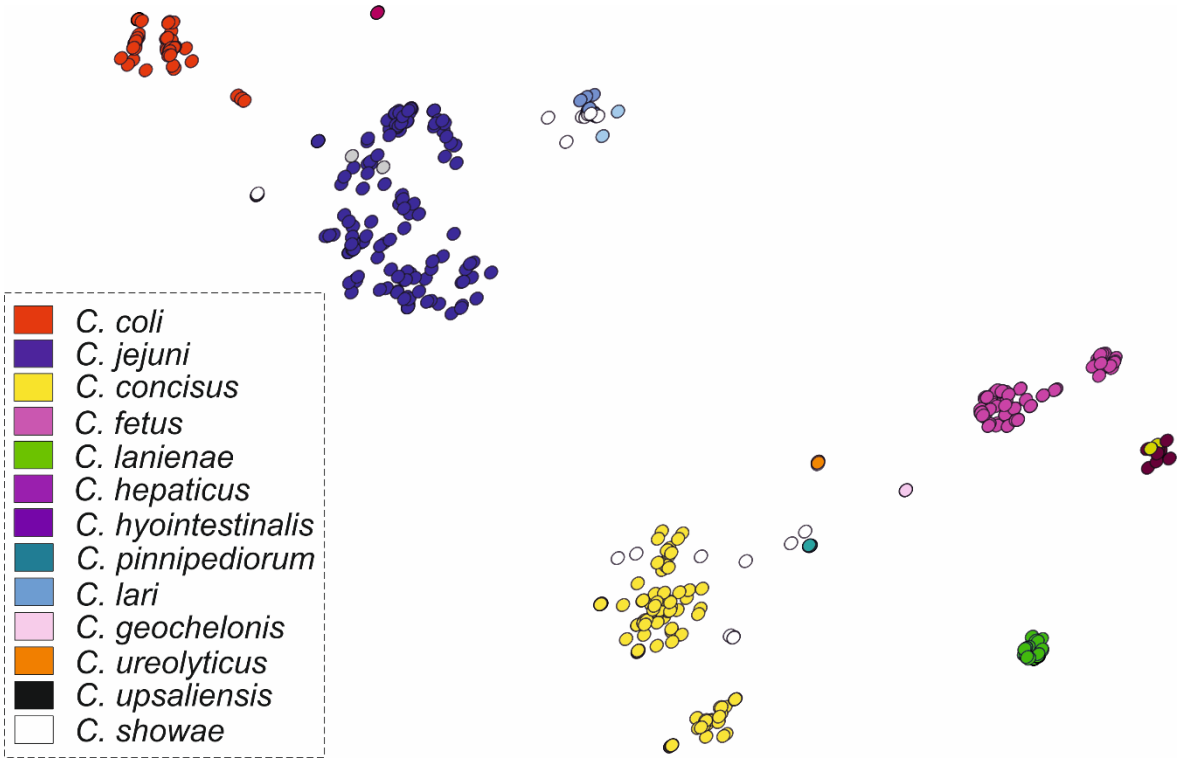


Figure 2.10. Accessory genome variation and the effect of HGT. Output of the PANINI algorithm for 631 isolates of the *Campylobacter* genus. The main species are coloured by different colours.

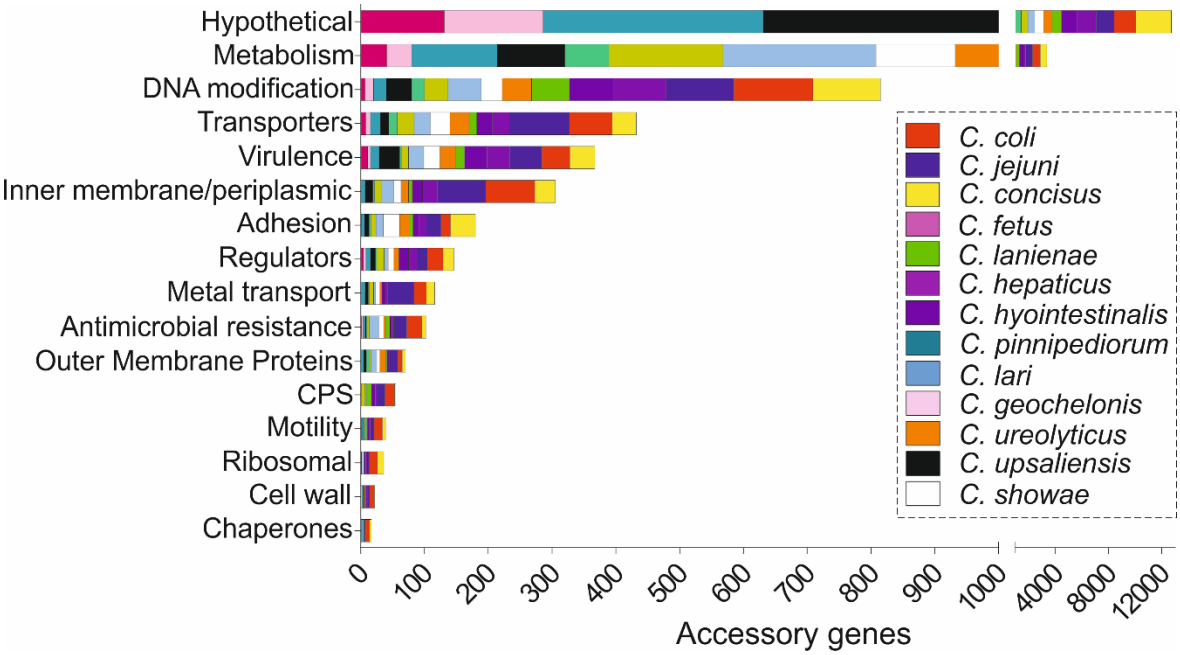


Figure 2.11. Accessory genome variation and the effect of HGT. Analysis of accessory gene function in all main *Campylobacter* species. The different gene functions are depicted on the *y* axis, while the number of shared accessory genes on the *x* axis.

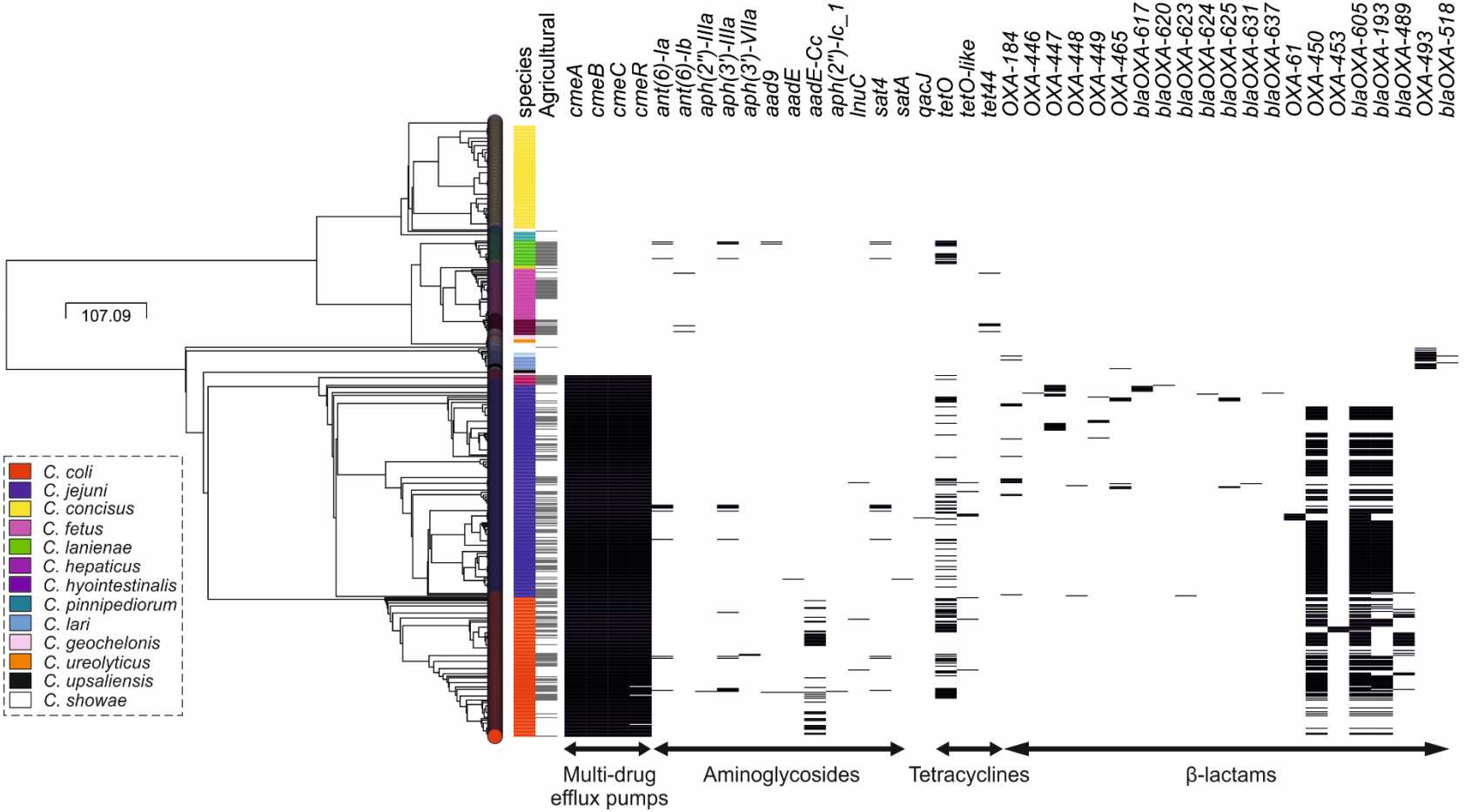


Figure 2.12. Accessory genome variation and the effect of HGT. Presence of antimicrobial resistance genes in the *Campylobacter* genus. The phylogenetic tree was reconstructed using a gene-by-gene concatenated alignment of 820 core genes and an approximation of the maximum-likelihood algorithm (ML) implemented in RAxML. The designated colour scheme was used for each species in the first column. The second column indicates whether the strain is isolated from an agricultural animal (grey). Remaining columns indicate presence of AMR genes (black). The scale represents the number of substitutions per site.

To investigate HGT further, work in this chapter focused on known examples of accessory genes that can move between different bacterial species (Mourkas et al., 2019; Sheinman et al., 2020). The genomes of all 631 isolates included in this chapter were compared to 8,762 known antibiotic resistance genes from the Comprehensive Antibiotic Resistance Database (CARD) (Jia et al., 2017), ResFinder (Zankari et al., 2012) and the National Center for Biotechnology Information (NCBI) databases. Homology (>75%) was found for 42 AMR determinants associated with multidrug efflux pumps, aminoglycosides, tetracyclines and β-lactams (Figure 2.12). Species that contained a significant proportion of agriculture-associated isolates, like *C. jejuni*, *C. coli*, *C. lanienae*, *C. hepaticus*, *C. hyointestinalis* and *C. fetus* contained far more AMR determinants (Figure 2.12). Twenty-seven and 18 AMR determinants from four antimicrobial drug classes, including tetracyclines, aminoglycosides, lincosamides and *β*-lactams, were found in *C. jejuni* and *C. coli*, respectively (Figure 2.12). Five and two AMR determinants, conferring resistance to aminoglycosides and tetracyclines, in *C. lanienae* and related *C. fetus* and *C. hyointestinalis*, respectively, were also detected (Figure 2.12). Tetracycline resistance *tetO* was detected in two *C. hepaticus* isolates (Figure 2.12). Genes associated with the cmeABC efflux pump system were highly prevalent in *C. jejuni*, *C. coli* and *C. hepaticus* (Figure 2.12). AMR determinants associated with *β*-lactams resistance, were present in marine-associated species *C. lari* (4), *C. subantarcticus* (1), *C. volucris* (1), *C. ornithocola* (3), *C. insulaenigrae* (1) and *Campylobacter sp.* (1) (Figure 2.12).

AMR genes are often found in Pathogenicity Islands (PAIs), in close proximity in the genome (Mourkas et al., 2019) and this analysis revealed several gene clusters that have been described in previous studies (Mourkas et al., 2019; Abril et al., 2010). The cluster of *ant(6)-Ia*, *sat-4*, *aph(3)-IIIa* genes in *C. jejuni* (chicken; n=5, clinical; n=1), *C. coli* (chicken; n=2, cattle; n=1) and *C. lanienae* (pig; n=1) isolates (Figure 2.12, Figure 2.13) has previously been described in *C. jejuni* and *C. coli* (Qin et al., 2012; Mourkas et al., 2019) but not in *C. lanienae*. The addition of *tetO* in this genetic association, was also observed in a clinical *C. jejuni* genome (Figure 2.12, Figure 2.13), consistent with previous findings (Mourkas et al., 2019). A cluster of *aad9*, *ant(6)-Ib* and *tetO* was detected in a *C. coli* isolate from water (Figure 2.12, Figure 2.13). The same genes, colocalized in a different ordered, were previously detected

in a *C. coli* isolates from sewage (Mourkas et al., 2019). Genes *aph(3)-IIIa*, *aad9*, *ant(6)-Ia* and *sat-4* were colocalized in two *C. lanienae* isolates from pigs (Figure 2.12, Figure 2.13).

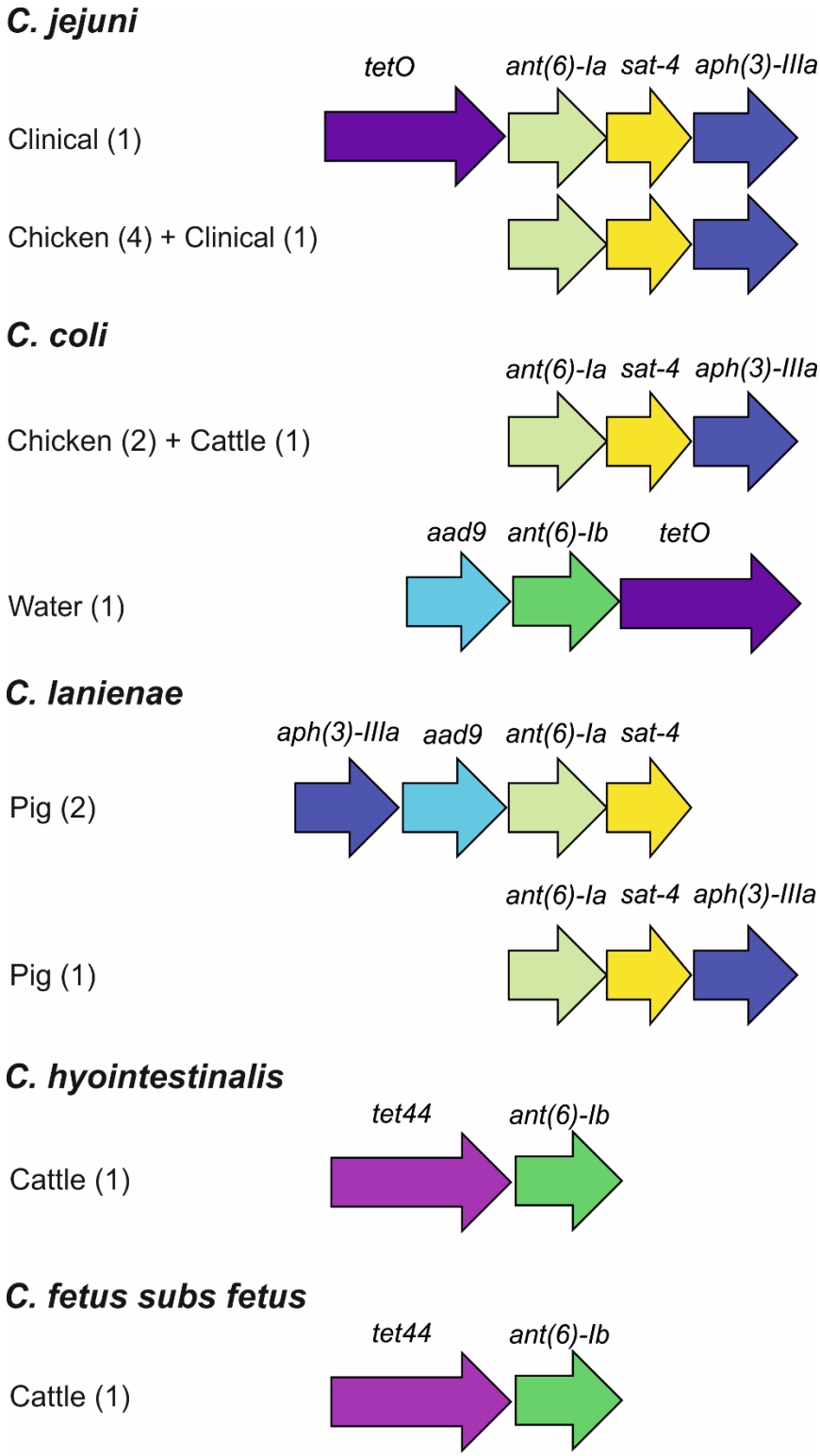


Figure 2.13. Genetic organization of AMR genes in *Campylobacter*. The presence of AMR genes, *tetO* (purple), *ant(6)-Ia* (olive green), *sat-4* (yellow), *aph(3)-IIIa* (blue), *aad9* (light blue), *ant(6)-Ib* (green) and *tet44* (fuchsia), is shown for representative genomes from *C. jejuni*, *C. coli*, *C. lanienae*, *C. hyointestinalis* and *C. fetus subspecies fetus* sampled from different agricultural animals. The number of isolate genomes containing each genomic arrangement is indicated in parenthesis.

These genes were located in a plasmid, on a genomic island along with other AMR genes, in a *C. coli* isolate from chicken (Mourkas et al., 2019). The last gene grouping included *tet44* and *ant(6)-Ia* genes, which were colocalized in *C. fetus subsp. fetus* (cattle; n=1) that has been previously described (Abril et al., 2010). This cluster (>75% sequence identity) was also detected in a *C. hyointestinalis* isolate from cattle (Figure 2.12, Figure 2.13). These findings are consistent with HGT-mediated circulation of AMR genes between different *Campylobacter* species and among host microbiome gene pools.

(2) Identifying adaptation in the core genome

Allelic variation in the core genome was investigated by calculating the number of core genome SNPs for all 15 major *Campylobacter* species (species with 3 or more isolates in the dataset). Major differences in the number of core genome SNPs ranging between 983 and 230,264, were found. *C. coli* had the largest range of SNPs (between 101,581 and 133,632) and had the highest mean number of SNPs together with *C. concisus* (131,279 and 227,395, respectively) (Figure 2.14), representing the most diverse species in the genus. *C. hepaticus* and *C. geochelonis* had the lowest mean number of SNPs (986 and 4,310, respectively), and despite the low number of isolates for these species, their values were lower compared to other species with equal number of isolates (Figure 2.14). Sampling bias could explain the low diversity of those species with isolates either being from hosts in close proximity (Piccirillo et al., 2016) or from contemporaneous outbreaks caused by genetically similar strains (Van et al., 2016).

Pairwise average nucleotide identities (ANI) for 820 genes, shared >95% by 605 isolates from 15 major species in the genus *Campylobacter*, were compared to all *C. jejuni* isolates. The analysis grouped species somewhat according to the genus population structure, and a two-stage similarity with *C. jejuni. C. coli* and *C. hepaticus* were the most similar to *C. jejuni*, followed by another grouping of *C. lari, C. subantarticus* and *C. upsaliensis* (Figure 2.15). Genus-wide genetic similarity and pairwise average nucleotide identity (ANI) for all possible pairs of genomes was calculated with FastANI (Jain et al., 2018). Isolates of the same bacterial species shared >95% similarity with each other (Figure 2.16), and signals of increased genetic similarities (between 85% and 90%) over greater phylogenetic distances

were observed throughout the matrix. These signals might represent potential genomic signatures (songs) between lineages (singers) that belong to different species that share the same host.

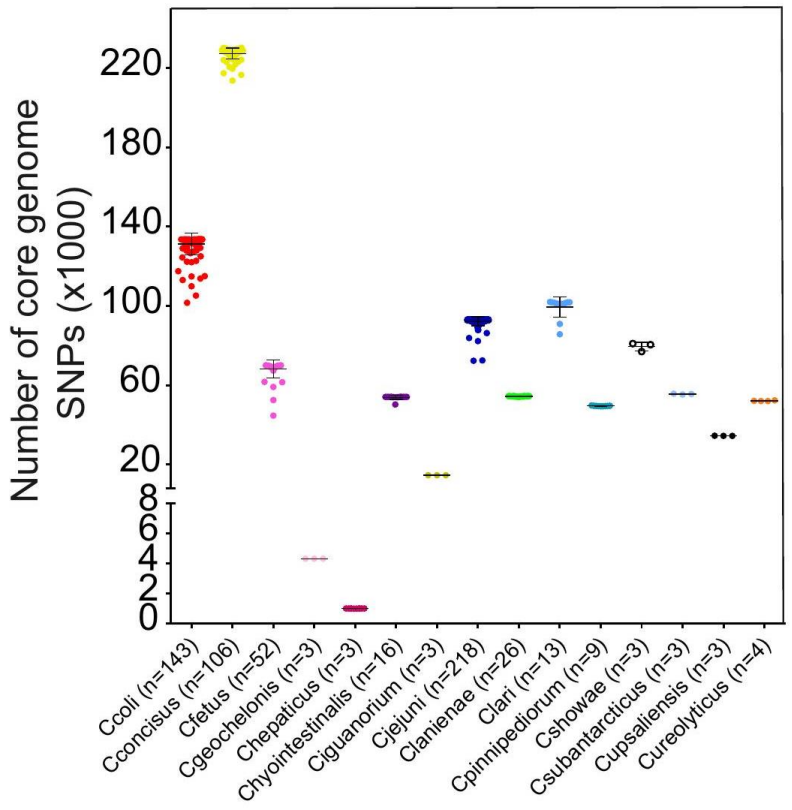


Figure 2.14. Core genome allelic variation. Number of SNPs per genome of the main *Campylobacter* species (where n≥3 isolates) in the core genome alignment. The horizontal line in each plot represents the mean value while the upper and lower lines the standard deviation.

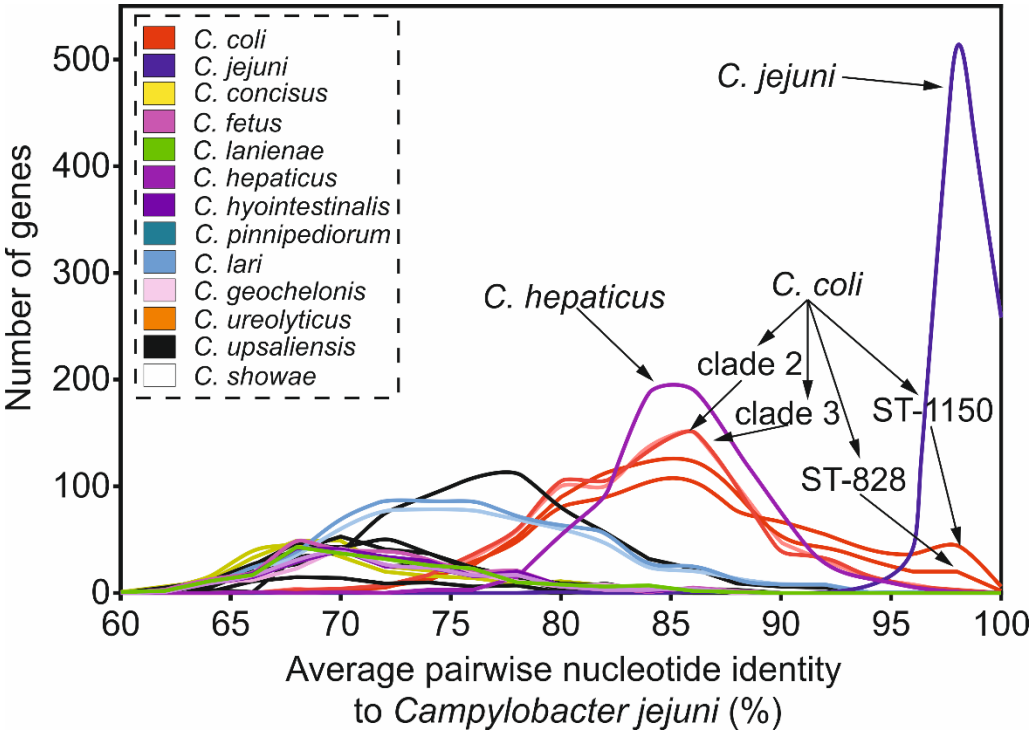


Figure 2.15. Average nucleotide identity for gene pairwise comparisons of 820 core genes for 605 genomes of 15 main *Campylobacter* species.

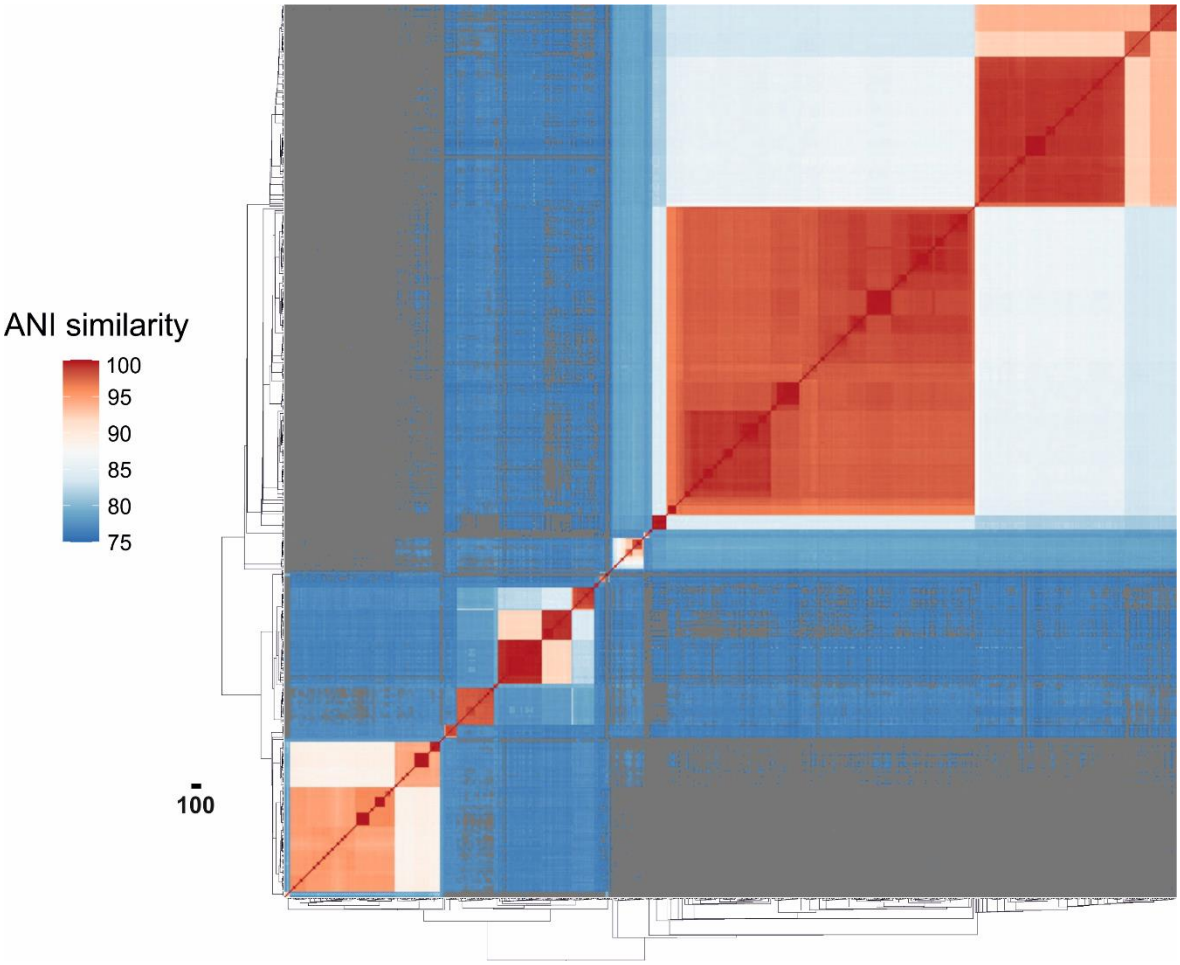


Figure 2.16. Pairwise average nucleotide identity comparison calculated for all 631 *Campylobacter* isolates shown on a heatmap with red indicating high (100%) and blue low (75%) nucleotide identity. Grey colour corresponds to missing ANI values <75% which are not calculated by FastANI (Jain et al., 2018).

(3) Enhanced within-host recombination

Intra-species recombination is common and has been reported in highly recombining bacteria, such as *Streptococcus suis* (Weinert et al., 2015), *Helicobacter pylori* (Berthenet et al., 2018), *Staphylococcus aureus* (Richardson et al., 2018; Murray et al., 2017), including *C. jejuni* (Mourkas et al., 2020), *C. coli* (Sheppard et al., 2010) and *C. fetus* (Iraola et al., 2017). Large scale recombination or introgression has also been described between *Salmonella* serovars (Criscuolo et al., 2019), and between *C. jejuni* and *C. coli,* generating hybrid lineages (Sheppard et al., 2008; Sheppard et al., 2011). To quantify genome-wide recombination in isolates cohabiting the same niche, Chromopainter/fineSTRUCTURE (Lawson et al., 2012) was used. The dataset was filtered, to include 297 isolates drawn from different species that share the same hosts or environments (Appendix, Table S2.1), and compared them in 27 donor-recipient groups. From a total number of 266,892 polymorphic sites, 258,444

(96.83%) recombining SNPs (mapped to 558 genes and representing 33.98% of the NCTC11168 genes) that had >90% probability of copying from a donor to a recipient strain, were detected. From those SNPs, 215,068 (83.21%), (mapped to 534 genes; 32.52% of the NCTC11168 genes) were found to be recombining within the same host. The hypothesis was supported in 10/27 group comparisons where more recombination was detected within than between host (Figure 2.17, Figure 2.18, Figure 2.19).

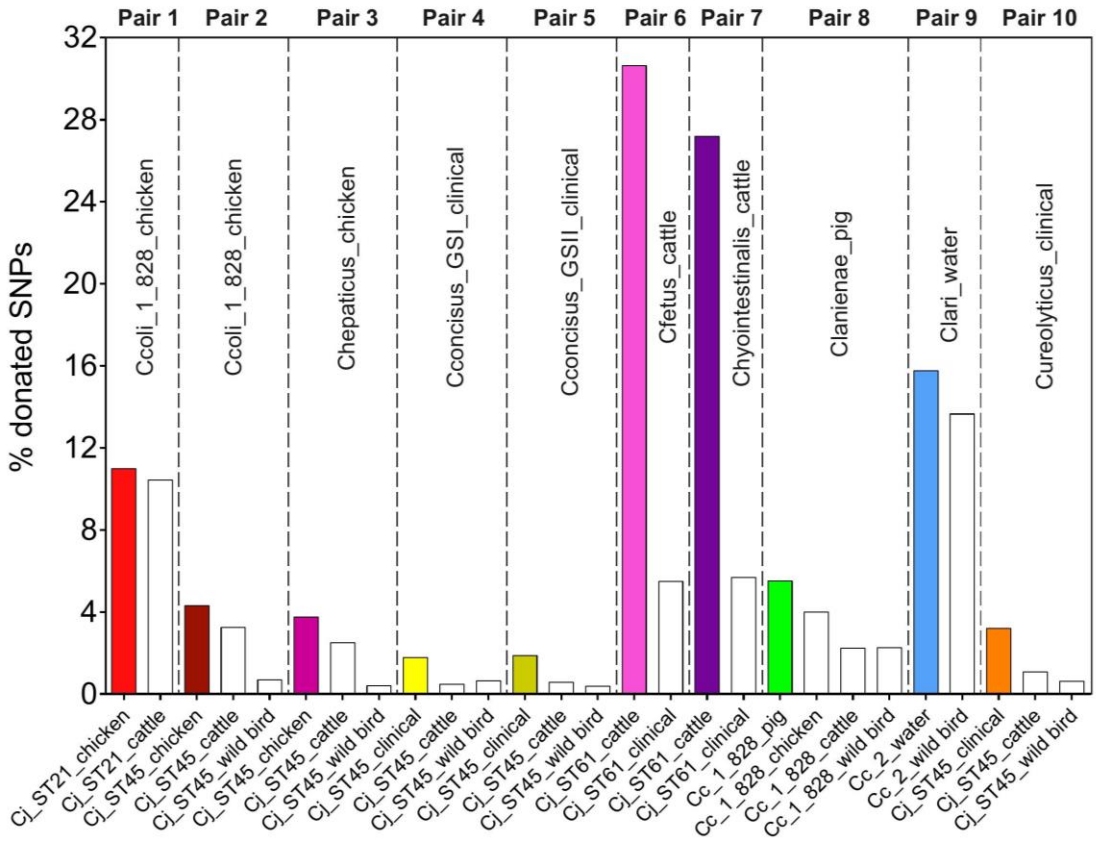


Figure 2.17. Recombination in donor-recipient comparisons. (A) The figure shows the recombining SNPs in 10 group donor-recipient comparisons. The proportion (%) of recombining SNPs with >90% probability of copying from a donor to a recipient genome is illustrated in the *y* axis. All donor groups are shows in the *x* axis. All coloured boxes correspond to comparison where donors and recipients are found in the same host.

To assess the robustness of the analysis and eliminate the likelihood of chance, the effect of randomization was included, and the analysis was repeated by assigning random hosts for every isolate (Appendix, Figure S2.1). In the 10 group comparisons that supported the hypothesis, 174,594 within-host recombining SNPs (mapped to 473 genes; 28.8% of NCTC11168 genes) and 109,564 between-host recombining SNPs (mapped to 395 genes; 24.05% of NCTC11168 genes), were detected. From the 473 within-host recombining genes, 45 genes contained the

highest number (>95th percentile) of recombining SNPs (Figure 2.18, Figure 2.19, Table S2.2). These genes had diverse inferred functions including metabolism, cell wall biogenesis, DNA modification, transcription and translation (Appendix, Table S2.2).

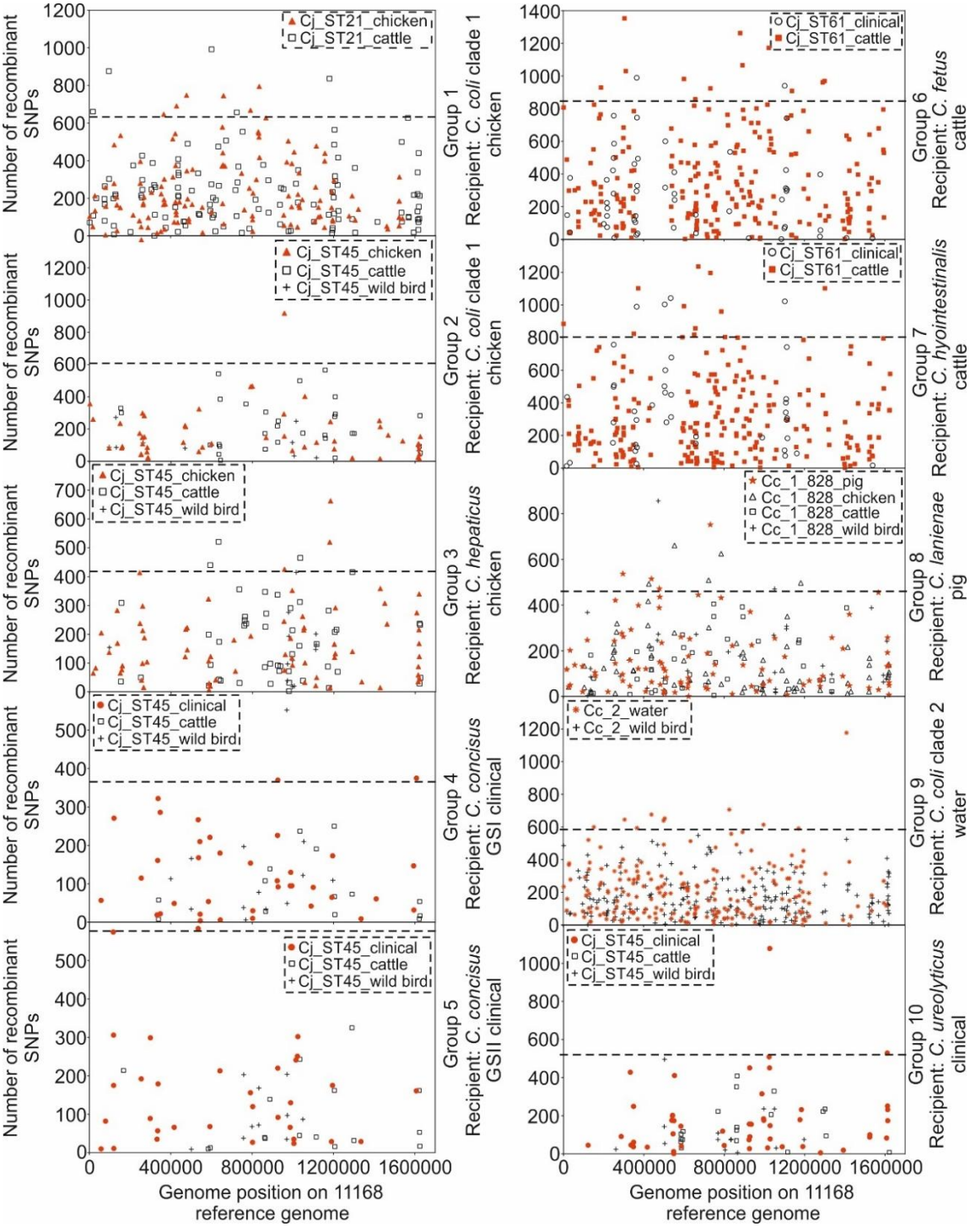


Figure 2.18. Genome position of genes containing recombining SNPs. Genes and their corresponding number of recombining SNPs, inferred by Chromosome Painting analysis for all ten group comparisons, and mapped to the NCTC11168 reference genome. Genes from within-host (red) and between-host (black, white) pair comparisons are shown for each group. Donor groups are isolates from chicken (triangle), cattle (square), wild bird (cross), pig (star), clinical (circle) and water

(snowflake) samples. The dashed line indicates the 95th percentile for every individual group comparison.

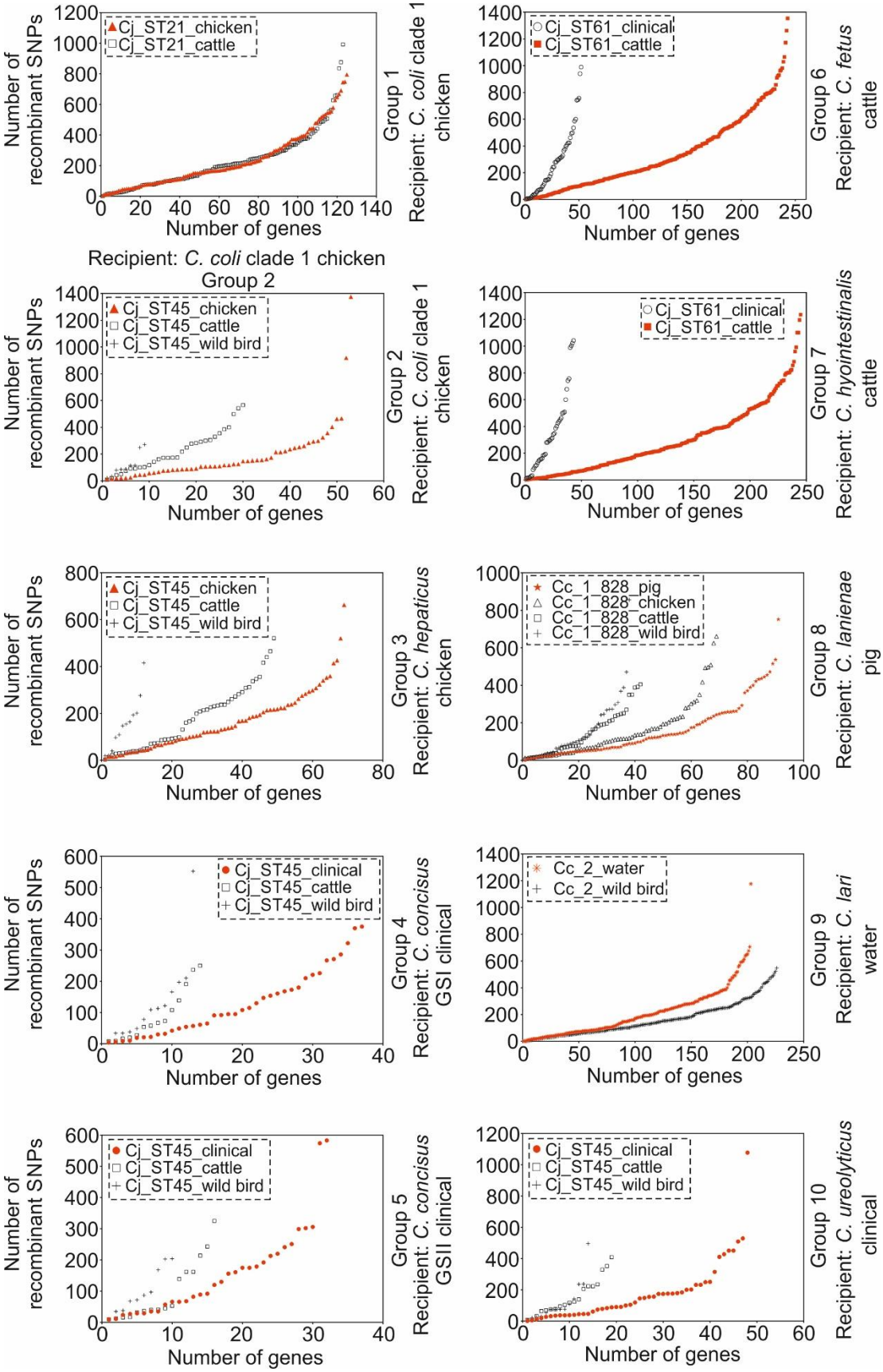


Figure 2.19. Genes ranked in ascending order of the number of recombining SNPs they contain as inferred by Chromosome Painting analysis for all ten group comparisons. Genes from within-host

(red) and between-host (black, white) are shown for each group. Donor groups are isolates from chicken (triangle), cattle (square), wild bird (cross), pig (star), clinical (circle) and water (snowflake) samples.

The within-host mobilome

Within-host interspecies recombination was observed for isolates sampled from chickens between generalist lineages ST-21 and ST-45 (donors; *C. jejuni*) and generalist ST-828 complex (recipient; *C. coli*). ClonalFrameML analysis detected a high recombination to mutation (*r*/*m*) for these lineages (Appendix, Table S2.3). Generally, agriculturally associated species/lineages had on average two times more *r*/*m* compared to non-agricultural associated ones (Appendix, Table S2.3). Genetic exchange between *C. jejuni* and *C. coli* has been described before (Sheppard et al., 2013). Analyses including *C. coli* ST-828 complex as a recipient, revealed *aroA* and *carB* genes as the most recombinogenic, with *C. jejuni* lineage ST-21 complex and ST-45 complex as donors, respectively (Appendix, Table S2.2). The first (*aroA*) encodes for an enzyme which is present in bacteria, yeast and plants and might be essential for growth and colonization in chickens (Wösten et al., 1996). The second one (*carB*) encodes for a carbamoylphosphate synthase and has been linked to biosynthesis of polysaccharides associated with virulence (McLennan et al., 2008). Generalist lineage ST-45 was involved as a donor in within-host interspecies comparisons with three more *Campylobacter* species, including *C. hepaticus* (chicken), *C. concisus* GSI and GSII (clinical) and *C. ureolyticus* (clinical) (Figure 2.17, Figure 2.18, Figure 2.19). This lineage had the highest *r*/*m* ratio from all other lineages or species involved in the comparisons (Appendix, Table S2.3). With *C. hepaticus* group as a recipient, the most recombining gene was *pnp* which encodes for a polynucleotide phosphorylase (Figure 2.20, Appendix, Table S2.2). Inactivation of this gene has been associated with decreased motility and lack of colonization in chickens (Haddad et al., 2012). For the clinical associated *C. concisus* the most highly recombining genes included *topA* and *ilvI* (Figure 2.20, Appendix, Table S2.2). While the first gene, encodes for a Topoisomerase Is that changes the DNA conformation during homologous recombination (Hyytiäinen et al., 2013), there is limited information for *ilvI* gene, involved in amino acid biosynthesis. For the emerging gastrointestinal pathogen *C. ureolyticus* comparison, the top recombining genes *secD* and *secY* are encoding for secretion proteins, which could be responsible for adaptation into a new host environment (Jiang and Fares, 2017).

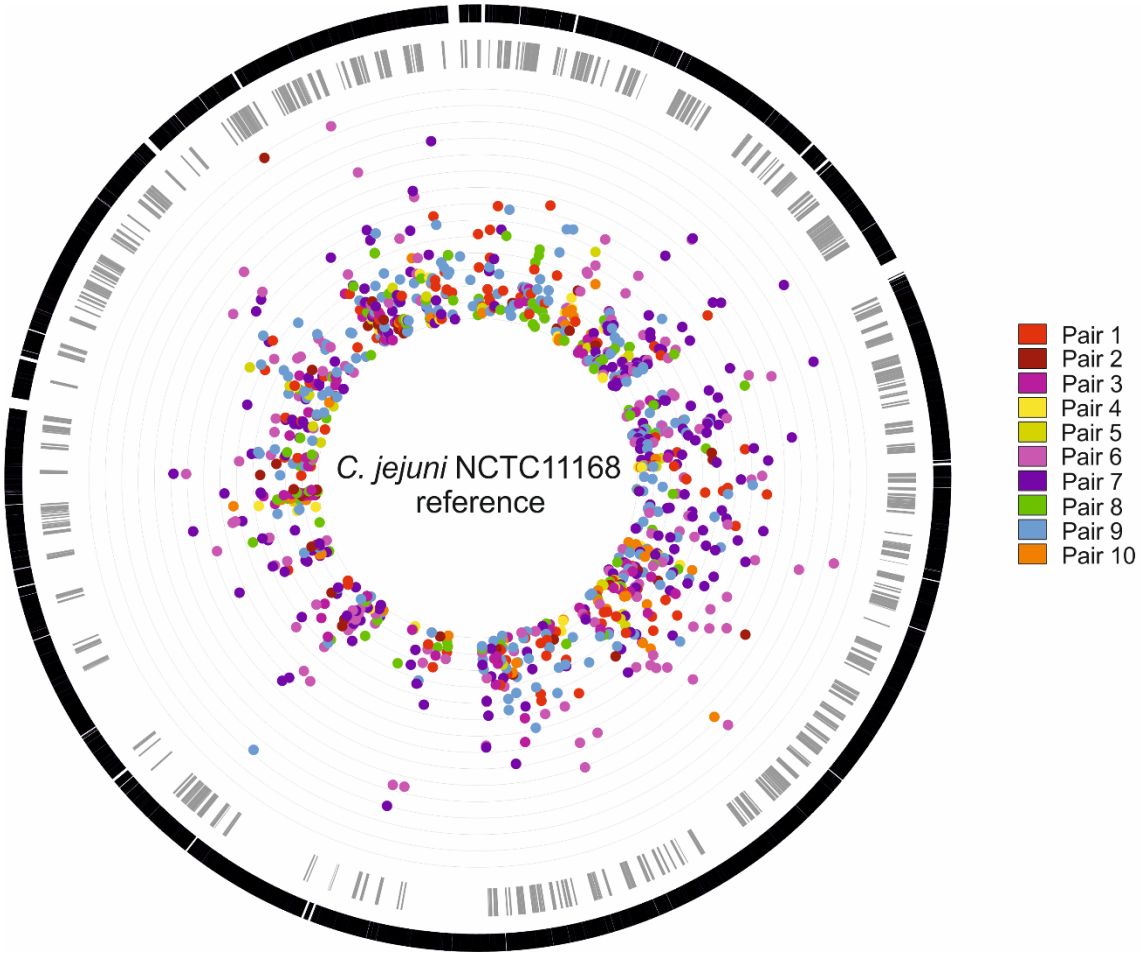


Figure 2.20. Pangenomic position of recombination analysis results as inferred by ChromosomePainting. Ticks in the outer circle (black) represent pangenomic positions of genes in the *C. jejuni* NCTC11168 reference genome used in this chapter. Ticks in the inner circle (grey) show the position of genes involved in within-host recombination. Genes containing recombining SNPs with >90% probability of copying from a donor to a recipient genome are represented along with a quantitative visualization of the number of the recombining SNPs detected in every gene for each group comparison (scale of quantification from 0 to 1375). Genes are coloured according to their corresponding group comparison.

There was increased recombination in genomes sampled from cattle between *C. jejuni* ST-61 lineage (donor; *C. jejuni*) and *C. fetus* and *C. hyointestinalis* (recipients) with 71.75% of all within-host recombining SNPs from all 10 groups detected in these two pair comparisons (Figure 2.17, Figure 2.18, Figure 2.19). Cattle-associated lineage ST-61 has previously been described as highly recombinogenic, and been associated with rapid clonal expansion and adaptation in cattle (Mourkas et al., 2020). The top recombining genes included *polA* and *dnaE* for *C. fetus* and *C. hyointestinalis* as recipients, respectively (Figure 2.17, Appendix, Table S2.2). These genes are involved in DNA replication/repair and homologous recombination

(Hyytiäinen et al., 2013). The group comparison with lineage ST-828 (*C. coli*) as donor and *C. lanienae* as a recipient showed increased recombination in pigs with *napA* being the top recombining gene (Figure 2.17, Appendix, Table S2.2). This gene is involved in electron transfer and is essential for nitrate reduction in *Campylobacter jejuni* (Pittman et al., 2007). Increased recombination was detected in isolates sampled from water involving environmental associated *C. coli* clade 2 as donor and *C. lari* as recipient (Figure 2.17). The top recombining gene in this analysis was *cj1476c* encoding for a pyruvate-flavodoxin oxidoreductase being differentially expressed under aerobic or microaerobic conditions (Gaynor et al., 2004).

Bacterial isolates inhabiting the same niche would likely require similar gene combinations to colonize and survive in the hew host environment. Recombination can promote the dissemination of adaptive genetic elements in isolates of different bacterial species. To identify host adaptive genes, the analysis focused on highly recombining genes (>95th percentile) in different group-species comparisons between strains that share the same host. Gene *gyrA* was found to be highly recombining in chickens (Groups 2 and 3) (Appendix, Table S2.2). Presence of AMR is common in livestock animals and a single mutation in the *gyrA* gene is known to confer resistance against ciprofloxacin (Luo et al., 2003). *C. jejuni* and *C. coli* isolates from clinical and animal samples are known to have high levels of resistance against ciprofloxacin (Food and Authority, 2019). In cattle, two genes, *aspS* and *glmS* were involved in recombination between three different *Campylobacter* species (Groups 5 and 6). The first is associated with virulence (Reddy and Zishiri, 2018) while the second is a conserved cell wall biosynthesis associated gene known to have multiple attachment transposon sites (Tn7) located downstream of its' genome position (Choi, 2009). Gene *napA* was highly recombining in cattle and pig hosts (Groups 5, 6 and 7) (Appendix, Table S2.2). This gene is essential for reducing the amounts of nitrate in the periplasm of gram-negative bacteria (Potter et al., 2001). These genes might represent candidate host adaptive genes conferring a strong advantage of colonization and survival in each animal host and environment.

The interspecies mobilome of the *Campylobacter* genus

Close contact between wildlife, livestock animals and humans can increase transmission opportunities between bacteria, where adaptive genes can be transferred between strains of different species, allowing for genetic variation while maintaining the species entities. However, parts of the genome were found to be recombining not only within, but also between hosts, indicating hotspots of recombination in the genome. Therefore, work on this chapter quantified the proportion of the mobile genome and focused at recombining SNPs found in more than one group comparison (Figure 2.21). These SNPs mapped to 337 genes (20.52% of the NCTC11168 genes; 2.15% of the pangenome) (Figure 2.22, Appendix, Table S2.4).

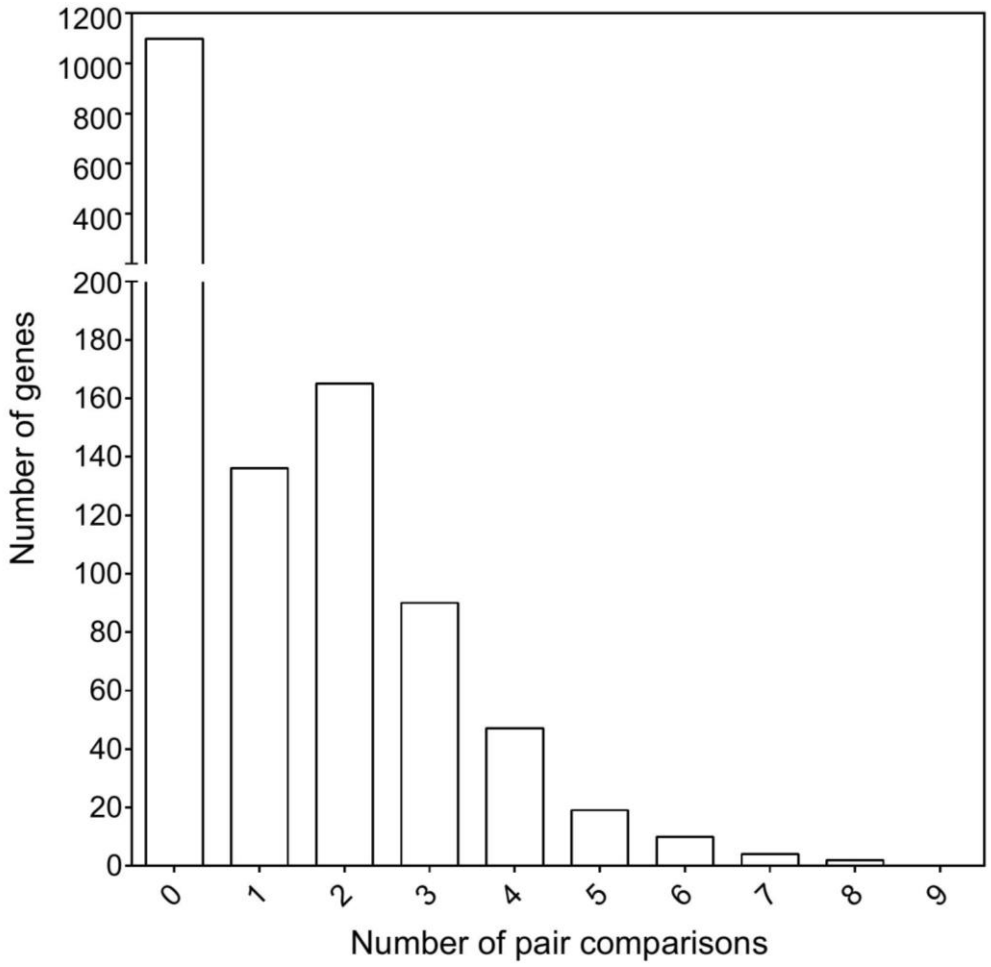


Figure 2.21. Genes recombining within the *Campylobacter* genus. The graph shows the number of genes involved in recombination in 10 different group comparisons. The number of genes is illustrated on the *y* axis while the increasing number of group comparisons on the *x* axis.

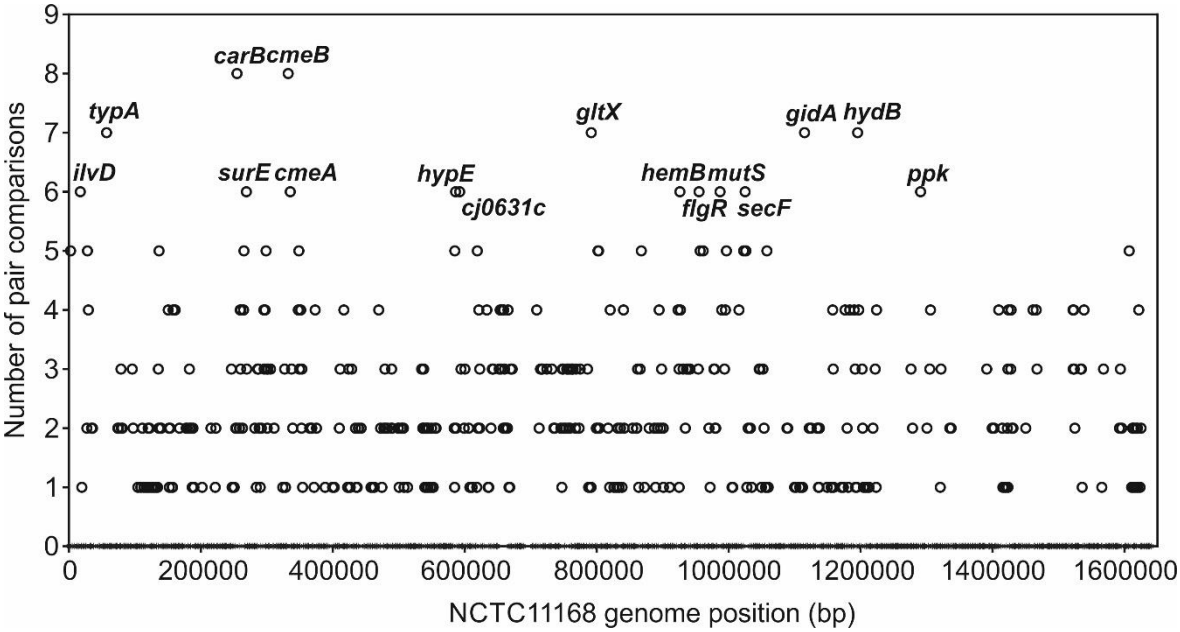


Figure 2.22. Genomic position of genes that are recombining within the *Campylobacter* genus. Presence of all recombining (o) and non-recombining (x) genes of the *C. jejuni* NCTC11168 reference genome. Genes present in at least six different group comparisons are depicted with their annotated name based on the *C. jejuni* NCTC11168 reference genome annotation.

To identify gene flow and analyse important predicted functions across many different *Campylobacter* species, analysis focused on 16 genes present in more than five group comparisons (Figure 2.22, Appendix, Table S2.4). Genes included *cmeA* and *cmeB* which are part of the predominant efflux pump CmeABC system in *Campylobacter.* Sequence variation in the drug-binding pocket of the *cmeB* gene has been linked to increased efflux function leading to resistance against multiple drugs (Yao et al., 2016). Use of antibiotics in human and veterinary medicine (outside of the EU) are interrelated with selection for AMR in *Campylobacter* species that are commonly isolated from livestock animals (Livermore, 2007). *Campylobacter* response to host transmission and virulence has been associated with biofilm formation and changes into surface polysaccharides (Szymanski et al., 2003; McLennan et al., 2008). Present in eight group comparisons, gene *carB*, encoding for a carbamoylphosphate synthase, has been associated with biosynthesis of substrates for many polysaccharides and is known to contain transposon insertion sites upstream of its' genomic location (McLennan et al., 2008). Other genes recombining in seven group comparisons included *typA*, *gltX*, *gidA* and *hydB*. While the function of the first two is not clear, *gidA* is associated with virulence in many bacteria (Mikheil et al., 2012) and *hydB* encodes for the large subunit of the H_2-uptake hydrogenase an essential enzyme involved in respiratory pathway in *C.*

concisus (Benoit and Maier, 2018). These genes might pinpoint at functions essential for bacteria to adapt into different hosts and environments.

Discussion

Current genomics methods focus on describing the genotypic diversity and population structure of bacteria by looking at gene variation in a pangenomic scale in large populations. It has become evident that there is movement of genetic material between lineages of the same or among multiple species occupying the same niche space. However, a rigorous approach of quantifying recombination in the absence of physical barriers, has yet to be applied. Through measurement of genetic exchange between species sharing a similar niche, work in this chapter provides evidence of cross-species recombination. This confounds traditional Linnaean approaches to Eukaryotic taxonomy and is consistent with previous commentary on difficulties assigning bacterial species (Shapiro et al., 2016; Mallet et al., 2016).

The *Campylobacter* genus revealed more structure and diversity compared to genera that contain species with traditionally low diversity, such as *Mycobacterium tuberculosis* (Gupta et al., 2018). In *Campylobacter*, distinct core genome clustering supports classification of distinct species clades, such as *C. jejuni* and *C. coli*. Evidence of introgression between agricultural associated *C. coli* and *C. jejuni* has been described before (Sheppard et al., 2013). Work in this chapter found evidence of subspecies clustering in *C. fetus* and genomospecies in *C. concisus* highlighting a division within those species. However, other species clades were not as genetically distinct, like the *C. lari* group marine clade including fuzzy species *C. subantarcticus, C. insulaenigrae, C. ornithocola, C. volucris, C. peloridis* and *C. lari*. Bacteria of the *Bacillus* genus also include species that do not form clear and distinct sequence clusters including *B. anthracis*, *B. cereus* and *B. thuringiensis* that are genetically very closely related (Helgason et al., 2000). Classifying different bacterial populations into discrete clusters proves to be challenging for some species where clear divisions do not exist. In *Campylobacter*, the presence of subspecies, hybrids and ‘fuzzy’ species appears to be the result of ecological factors affecting the population structure in each species.

Exploring the source of *Campylobacter* species, revealed a broad diversity of hosts and environments (Figure 2.3). Notably, many *Campylobacter* species were isolated from livestock animals (Figure 2.5), probably due to the expanding niche driven by animal domestication and intensive livestock production over the last 100 years. Most uncommon species were usually isolated from one source (reptiles, marine mammals, wild birds) (Figure 2.5), revealing ecology driven adaptation in a particular niche. For example, isolates from *C. fetus* subsp *testudinum*, *C. iguanorium* and *C. geochelonis* were sampled from reptile-related animals, while isolates from *C. pinnipediorum* were sampled from seals, indicating adaptation to a single wild animal host species. Similarly, *C. jejuni* and *C. coli* isolates have been sampled from various wild bird species (Griekspoor et al., 2013; Cody et al., 2015). These wild animal associated *Campylobacters* are occasionally linked with human disease, indicative of infrequent contact with humans which decreases the risk of spillover.

Genetic variation was observed at the number of core and accessory genes between species. For some species, like *C. hepaticus*, this could be the result of sampling bias that mirror low genetic diversity (outbreaks). A large proportion of species-specific accessory genes encoding hypothetical proteins highlights a need for further gene function characterization. Nevertheless, all species shared a core set of genes important for main metabolic functions. Genome-wide introgression between two different bacterial strains or species can result in the generation of hybrid strains (Sheppard et al., 2013; Criscuolo et al., 2019). Investigating the core genome allelic variation of every species in the genus provides clues about possible formation of hybrid groups (Mourkas et al., 2020; Sheppard et al., 2013). While this is true for the agricultural associated *C. coli* and *C. jejuni* (Sheppard et al., 2013), there was no evidence of genetic admixture in any of the other *Campylobacter* species as a result of hybridization. This could imply that recombination does not happen randomly, and that hybridization is not common in distantly related *Campylobacter*. The measure of ANI suggests that strains that belong to the same species share roughly >95% nucleotide identity (Jain et al., 2018). Signatures of genetic similarities between strains of different *Campylobacter* species (Figure 2.16) may be translated as flow of genetic elements essential for inhabiting a niche.

Exchange of large DNA fragments by horizontal gene transfer (HGT) is common between bacterial species (Tran and Boedicker, 2017). Antimicrobials exert a strong selective pressure, bacteria can overcome either by point mutations in genes of interest (Lapierre et al., 2016), or by acquiring genes that confer resistance against antibiotics through HGT (Yahara et al., 2014; Yahara et al., 2016). Studies suggest that HGT is a recurrent process, with AMR genes being widely transferred, not only between species of the same genus (Baker et al., 2018; Ochman et al., 2000), but also across different genera of bacteria (Mourkas et al., 2019; Park and Andam, 2020; Eisen, 2000). Every host and environment represent a bacterial niche, comprised of micro-niches, subjected to different selection pressures. The ecological and host colonization profile of different *Campylobacter* species provides scenarios where physical barriers to recombination are lifted, in cases of closely related species occupying the same niche (Figure 2.9). This grants the unique opportunity to quantify and compare within and between-niche recombination.

Evidence for HGT has been demonstrated for AMR genes among isolates from livestock, human and sewage and between isolates of two different *Campylobacter* species (*C. jejuni* and *C. coli*) (Mourkas et al., 2019). Work in this chapter identified AMR genes present in strains from six agricultural associated species including *C. jejuni*, *C. coli*, *C. hepaticus*, *C. lanienae*, *C. fetus* and *C. hyointestinalis* (Figure 2.12). This is not surprising considering the recently excessive use of antibiotics in veterinary medicine which allows farm animals to act as hosts for gene pools of AMR genes (Teuber, 2001). Presence of AMR genes was also found in six species of the *C. lari* group, including *C. lari*, *C. subantarcticus*, *C. volucris*, *C. insualenigrae*, *C. ornithocola* and *Campylobacter spp.* all of which associated with marine mammals, wild birds and environmental water (Figure 2.12). Various pathogenic bacteria, including *Salmonella* (Olsen et al., 1996), *Campylobacter* (Griekspoor et al., 2009; Broman et al., 2000) and *Escherichia coli* (Hernández et al., 2012) have been isolated from wild birds and sea waters, many of which being resistance to different antibiotics, indicating AMR transmission through wild bird migration and possibly spread into other species via HGT (Aminov and Mackie, 2007; Hernández and González-Acuña, 2016). Transfer of genomic islands containing AMR gene clusters can be facilitated via natural transformation or plasmid mobilization in a specific niche or environment. Analysis in this chapter revealed AMR gene clusters in *Campylobacter* genomes associated with different farm animals. In some cases,

strains from phylogenetically closely related species (*C. fetus* and *C. hyointestinalis*) isolated from cattle, shared the same AMR gene cluster (*tet44* and *ant(6)-Ib*) described before in *C. fetus* subsp. *fetus* (Abril et al., 2010). The presence or absence of AMR genes could be the result of genes that share the shame ancestry, however the syntenic rearrangement of AMR gene blocks indicates the circulation of colocalized AMR genes among related species and host niche gene pools and is consistent with HGT.

For DNA exchange to take place, bacterial strains need to be in close proximity. For *Campylobacter* strains that would most likely be the gut of an animal. *Campylobacter* species undergo high levels of homologous and non-homologous recombination which can vary among lineages and across species (Yahara et al., 2016; Mourkas et al., 2020; Iraola et al., 2017; Wilson et al., 2009; Sheppard et al., 2014). Characterizing the segments of DNA that are involved in recombination provides information about gene function and the proportion of the mobile genome (Yahara et al., 2014). To identify this, analysis focused on quantifying the DNA that is transferred between strains from different species that may or may not be found in the same host. The analysis revealed specific donor-recipient pairs with enhanced recombination within rather than between a certain niche (Figure 2.17). Many group comparisons failed to support the hypothesis. Although the exact reasons are not clear, mechanistic barriers could prevent recombination for some species that share the same niche. Furthermore, physical isolation of species occupying different subniches in every host could be the reason. For example in poultry, *C. hepaticus* is commonly isolated from the liver (Van et al., 2016), while *Campylobacter* strains from other species (*C. jejuni*, *C. coli*) are commonly isolated from different parts of the gastrointestinal tract. Strains from the generalist *C. jejuni* ST-45 were involved in half of the group comparisons that supported the hypothesis (Figure 2.17). Given their generalist lifestyle of adapting in various niches and their high recombinogenic nature, isolates of this lineage might be able to colonize different subniches within hosts and environments. The host attribution signal is commonly lost in generalist lineages, including *C. jejuni* ST-21, ST-45 complexes and *C. coli* ST-828 complex, due to recent rapid host switches (Dearlove et al., 2016). However, signatures of adaptation have been observed between strains isolated from cattle compared to chickens, in ST-45 complex (Sheppard et al., 2013). Extended sampling from

various subniches in different animals is needed to further understand the impact of within-host microecology.

Every species group comparison revealed genomic hot spots of recombination throughout the genome, involving genes with putative functions related to virulence, colonization, DNA modification or adaptation into a new niche. Different parts of the genome evolve and speciate at different rates. In fact, some genes might not speciate, but rather coevolve with the host forming a host adapted gene pool. Identifying highly recombinant genes in different species could provide information about host-specific gene pools. Donor-recipient analyses revealed highly recombinogenic genes from multiple group-species comparisons which shared the same host. *Campylobacter* isolates from chickens are highly resistance to fluoroquinolones, and particularly ciprofloxacin, which is the drug of choice to treat clinical campylobacteriosis cases, with mutations in the *gyrA* gene conferring resistance (Hormeño et al., 2016). Therefore, selective pressure from this antibiotic may explain the detection of this gene as highly recombinant in chicken isolates. Group-species comparisons in cattle showed very high levels of recombination (2/3 of all recombining SNPs). The highly recombinogenic nature of *C. jejuni* ST-61 complex, involved in these comparisons as a donor, could explain the high amount of recombination compared to the other species-group comparisons. Transposons are mobile genetic elements that can jump from one genome position to another between strains. Conjugative transposons integrate in a specific site (Tn7) which is located downstream of *glmS* gene (Choi, 2009). Analogs of *glmS* are found in many bacteria with multiple Tn7 sites downstream of that gene (Choi, 2009) which could explain the high recombination of this genomic region in cattle. To grow under microaerobic conditions, *Campylobacter* is equipped with *napA*, a gene encoding a nitrate reductase (Pittman et al., 2007). Nitrate, if converted to nitrite, can cause food poisoning to farm animals. Accumulation of high amounts of nitrate in plants, slurry, straw and drainage water could be toxic to livestock, including cattle and pigs (Alexander et al., 2009), and thus an actively expressed or beneficial allele of *napA* gene might be needed in strains colonizing the gut of those animals. While speculations about how the putative function of those genes can promote adaptation into different hosts, further *in vitro* and *in vivo* experiments are needed to confirm these assumptions.

Recombination allows for rapid adaptation into a host environment favouring adaptive genes to spread through bacteria populations. Work in this chapter found efflux pump genes *cmeA* and *cmeB*, associated with MDR resistance, to be recombining in multiple donor-recipient group-species, possibly driven by the presence of antibiotics in the surrounding environment. Most prokaryotes contain a *carAB* operon which encodes for carbamoylphosphatase an enzyme associated with polysaccharide biosynthesis potentially linked to virulence and better growth (McLennan et al., 2008). Gene *carB*, encoding for the large subunit of this enzyme, was recombining in all the group-species comparisons highlighting a possibly essential role of this gene for adaptation in different environments. Hydrogen is an electron donor to bacterial electron transport chains and is abundant in the gut of birds and mammals. *Campylobacter jejuni* encodes a NiFe uptake hydrogenase, which is activated by the presence of nickel in the environment (Howlett et al., 2012). Gene *hydB*, encoding the NiFe active site in the large subunit of this enzyme, might confer an advantage into utilizing nickel in a new niche, which could explain why it is recombining in multiple species-group comparisons. These genes reveal information about potentially essential functions in different hosts and environments and provide a basis for further *in vitro* experiments to investigate their role in bacterial adaptation.

The work in this chapter quantifies and compares the rate of recombination within-species and within-niches (specifically agriculture). Genes that can spread (likely via HGT) across species boundaries, were identified. Subsequent changes in bacterial phenotypes make speciation difficult and support theories promoting gene-centric descriptions of bacterial evolution. Analysis detected 0.6 times more recombination within- rather than between-host for some species-groups, revealing ecology as a driving factor for niche adapted genes. As different lineages are inhabiting a niche, gene ecology plays an important role in maintaining a metabolic network across bacterial species. Humans continue to change the environmental landscape creating novel niches for bacteria. By understanding the mechanisms of evolution, scientists can hope to decrease the risk of these versatile bacterial pathogens to spread into human populations.

Materials & Methods

Genome sequence collection and archiving

A total of 631 *Campylobacter*, 18 *Arcobacter*, eight *Sulfurospirillum* and five *Helicobacter* genomes were assembled from the NCBI database and previously published genomes (Appendix, Table S2.1). Isolates were sampled from clinical cases of campylobacteriosis and faeces from chickens, ruminants, wild birds, wild mammals, pets and environmental sources. Genomes and related metadata were uploaded and archived into BIGS database (Sheppard et al., 2012).

Sequence isolates

Comparative genomics analyses focused on all 631 *Campylobacter* species available on public repositories and identified until the conception of this chapter. The dataset consists of *C. avium* (n=1), *C. coli* (n=143), *C. concisus* (n=106), *C. corcagiensis* (n=1), *C. cuniculorum* (n=2), *C. curvus* (n=2), *C. fetus* (n=52), *C. geochelonis* (n=3), *C. gracilis* (n=2), *C. helveticus* (n=1), *C. hepaticus* (n=10), *C. hominis* (n=1), *C. hyointestinalis* (n=16), *C. iguanorium* (n=3), *C. insulaenigrae* (n=1), *C. jejuni* (n=218), *C. lanienae* (n=26), *C. lari* (n=13), *C. mucosalis* (n=1), *C. ornithocola* (n=1), *C. peloridis* (n=1), *C. pinnipediorum* (n=9), *C. rectus* (n=1), *C. showae* (n=3), *C. sputorum* (n=1), *C. subantarcticus* (n=3), *C. upsaliensis* (n=3), *C. ureolyticus* (n=4), *C. volucris* (n=2) and *Campylobacter sp* (n=1) (Appendix, Table S2.1).

Pangenome creation, core and accessory gene variation, phylogenetic analysis

Sequence data were analysed using PIRATE, a fast and scalable pangenomics tool which allows for orthologue gene clustering in divergent bacterial species (Bayliss et al., 2019). Genomes were annotated in Prokka (Seemann, 2014), using a genus database consisting of well annotated *C. jejuni* strains NCTC11168, 81116, 81-176, M1 and plasmids pTet and pVir in addition to the already existing databases used by Prokka (Seemann, 2014). Briefly, annotated genomes were used as input for PIRATE. Non-redundant representative sequences were produced using CD-HIT and the longest sequence was used as a reference for sequence similarity interrogation using BLAST/DIAMOND. Gene orthologues were defined as “gene families” and were clustered in different MCL thresholds, from 10 to 98 % sequence identity (10, 20, 30, 40, 50, 60, 70, 80, 90, 95, 98). Higher MCL thresholds were used to identify allelic variation within different loci. An inflation value of 4 was used to increase the granularity of MCL clustering between gene families. BLAST high-

scoring pairs (HPSs) with a reciprocal minimum length of 90% of the query/subject sequence were excluded from MCL clustering in order to reduce the number of spurious associations between distantly related or conserved genes. This information was used to generate gene presence/absence and allelic variation matrices. A core gene-by-gene multiple sequence alignment was generated using MAFFT (Katoh, 2002) containing genes shared >95% by all isolates. Phylogenetic trees, based on core gene-by-gene alignments, were reconstructed using the maximum-likelihood algorithm implemented in RAxML v8.2.11 (Stamatakis, 2014) with GTRGAMMA as substitution model. [N.B. The PIRATE analysis was performed in collaboration with Sion Bayliss, University of Bath, Bath, UK].

Quantifying accessory and core genome variation

To examine whether the accessory gene content of all 631 *Campylobacter* genomes is consistent with the core gene one, the PANINI software, which visualizes the output as a two-dimensional projection (Abudahab et al., 2019), was used. Subsequently, all *Campylobacter* genomes were screened for the presence of antimicrobial resistance genes against the CARD (Jia et al., 2017), ResFinder (Zankari et al., 2012) and NCBI databases. A positive hit was defined when a gene had >75% nucleotide identity over >50% of the sequence length. A gene presence/absence matrix for every antimicrobial resistance gene was generated for every genome. Genomes carrying AMR genes were screened to characterize the location of adjacent genes using SnapGene software (GSL Biotech; available at snapgene.com), as previously described (Mourkas et al., 2019).The number of core SNPs was identified using SNP-sites (v2.3.2) (Page et al., 2016). The program FastANI v.1.0 (Jain et al., 2018) was used to analyse the degree of genomic relatedness and similarity of all 631 *Campylobacter* genomes. This software provides a robust way to examine intra and interspecies genetic relationships between bacterial genomes. This generated a lower triangular matrix with the lowest ANI value at 75% (as computed by FastANI).

Inference of recombination regions

Whole genome sequences of 297 *Campylobacter* strains, isolated from different hosts or environments, were selected to infer co-ancestry using Chromopainter/fineSTRUCTURE (Lawson et al., 2012) as previously described (Pascoe et al., 2017; Yahara et al., 2018). This resulted in 27 donor-recipient group

comparisons, representing all possible combinations where strains from different species were isolated from the same host. To get statistical power, donor groups that contained at least eight isolates from the same source, were selected. *C. jejuni* and *C. coli* clade 1 genomes isolated from seals and water were excluded from the analysis, as these most likely represent spillover events and not true host segregated genomes. Briefly, a pairwise genome alignment between reference genome NCTC11168 and one of the 297 strains was generated using progressiveMauve (Darling et al., 2010). This enabled the construction of positional homology alignments for all genomes regardless gene content and genome rearrangements which were then combined into a multiple whole-genome alignment, as described before (Yahara et al., 2018). Chromopainter/fineSTRUCTURE was then used to calculate the amount of DNA sequence that is donated between a donor and a recipient (Lawson et al., 2012). ClonalFrameML (Didelot and Wilson, 2015) was used to infer the relative number of substitutions introduced by recombination (*r*) and mutation (*m*) as the ratio *r*/*m*, as described before (Mourkas et al., 2020). [N.B. The Chromopainter /fineSTRUCTURE analysis was performed in collaboration with Koji Yahara, Antimicrobial Resistance Center, National Institute of Infectious diseases, Tokyo, Japan].

References

Abril, C., Brodard, I. and Perreten, V., 2010. Two Novel Antibiotic Resistance Genes, tet(44) and ant(6)-Ib, Are Located within a Transferable Pathogenicity Island in Campylobacter fetus subsp. fetus. *Antimicrobial Agents and Chemotherapy*, 54(7), pp.3052–3055.

Abudahab, K., Prada, J.M., Yang, Z., Bentley, S.D., Croucher, N.J., Corander, J. and Aanensen, D.M., 2019. PANINI: Pangenome Neighbour Identification for Bacterial Populations. *Microbial Genomics*, 5(4).

Alexander, J., Benford, D., Cockburn, A., Cravedi, J., Dogliotti, E., Domenico, A. Di, Férnandez-cruz, M.L., Fürst, P., Fink-gremmels, J., Galli, C.L., Grandjean, P., Gzyl, J., Heinemeyer, G., Mutti, A., Schlatter, J., Leeuwen, R. Van and Peteghem, C. Van, 2009. Nitrite as undesirable substances in animal feed - Scientific Opinion of the Panel on Contaminants in the Food Chain. *EFSA Journal*, 7(4), pp.1–47.

Aminov, R.I. and Mackie, R.I., 2007. Evolution and ecology of antibiotic resistance genes. *FEMS Microbiology Letters*, 271(2), pp.147–161.

Baker, K.S., Dallman, T.J., Field, N., Childs, T., Mitchell, H., Day, M., Weill, F.-X., Lefèvre, S., Tourdjman, M., Hughes, G., Jenkins, C. and Thomson, N., 2018. Horizontal antimicrobial resistance transfer drives epidemics of multiple Shigella species. *Nature communications*, 9(1), p.1462.

Bayliss, S.C., Thorpe, H.A., Coyle, N.M., Sheppard, S.K. and Feil, E.J., 2019. PIRATE: A fast and scalable pangenomics toolbox for clustering diverged orthologues in bacteria. *GigaScience*, 8(10), pp.1–9.

Benoit, S.L. and Maier, R.J., 2018. Site-directed mutagenesis of Campylobacter concisus respiratory genes provides insight into the pathogen's growth requirements. *Scientific Reports*, 8(1), p.14203.

Berthenet, E., Yahara, K., Thorell, K., Pascoe, B., Meric, G., Mikhail, J.M., Engstrand, L., Enroth, H., Burette, A., Megraud, F., Varon, C., Atherton, J.C., Smith, S., Wilkinson, T.S., Hitchings, M.D., Falush, D. and Sheppard, S.K., 2018. A GWAS on Helicobacter pylori strains points to genetic variants associated with gastric cancer risk. *BMC Biology*, 16(1), p.84.

Broman, T., Bergström, S., On, S.L.W., Palmgren, H., McCafferty, D.J., Sellin, M. and Olsen, B., 2000. Isolation and Characterization of Campylobacter jejuni subsp. jejuni from Macaroni Penguins (Eudyptes chrysolophus) in the Subantarctic Region. *Applied and Environmental Microbiology*, 66(1), pp.449–452.

Choi, K.-H., 2009. Applications of Transposon-Based Gene Delivery System in Bacteria. *Journal of Microbiology and Biotechnology*, 19(3), pp.217–28.

Cody, A.J., McCarthy, N.D., Bray, J.E., Wimalarathna, H.M.L., Colles, F.M., Jansen van Rensburg, M.J., Dingle, K.E., Waldenström, J. and Maiden, M.C.J., 2015. Wild bird-associated C ampylobacter jejuni isolates are a consistent source of human disease, in Oxfordshire, United Kingdom. *Environmental Microbiology Reports*, 7(5), pp.782–788.

Colles, F.M., Dingle, K.E., Cody, A.J. and Maiden, M.C.J., 2008. Comparison of Campylobacter Populations in Wild Geese with Those in Starlings and Free-Range Poultry on the Same Farm. *Applied and Environmental Microbiology*,

74(11), pp.3583–3590.

Colles, F.M., McCarthy, N.D., Bliss, C.M., Layton, R. and Maiden, M.C.J., 2015. The long-term dynamics of C ampylobacter colonizing a free-range broiler breeder flock: an observational study. *Environmental Microbiology*, 17(4), pp.938–946.

Criscuolo, A., Issenhuth-Jeanjean, S., Didelot, X., Thorell, K., Hale, J., Parkhill, J., Thomson, N.R., Weill, F.-X., Falush, D. and Brisse, S., 2019. The speciation and hybridization history of the genus Salmonella. *Microbial Genomics*, 5(8), pp.1–11.

Darling, A.E., Mau, B. and Perna, N.T., 2010. progressiveMauve: Multiple Genome Alignment with Gene Gain, Loss and Rearrangement. *PLoS ONE*, 5(6), p.e11147.

Davenport, E.R., Sanders, J.G., Song, S.J., Amato, K.R., Clark, A.G. and Knight, R., 2017. The human microbiome in evolution. *BMC Biology*, 15(1), p.127.

Dearlove, B.L., Cody, A.J., Pascoe, B., Méric, G., Wilson, D.J. and Sheppard, S.K., 2016. Rapid host switching in generalist Campylobacter strains erodes the signal for tracing human infections. *The ISME Journal*, 10(3), pp.721–729.

Debruyne, L., On, S.L.W., De Brandt, E. and Vandamme, P., 2009. Novel Campylobacter lari-like bacteria from humans and molluscs: description of Campylobacter peloridis sp. nov., Campylobacter lari subsp. concheus subsp. nov. and Campylobacter lari subsp. lari subsp. nov. *International journal of systematic and evolutionary microbiology*, 59(Pt 5), pp.1126–32.

Didelot, X. and Wilson, D.J., 2015. ClonalFrameML: Efficient Inference of Recombination in Whole Bacterial Genomes. *PLOS Computational Biology*, 11(2), p.e1004041.

Dingle, K.E., Colles, F.M., Wareing, D.R., Ure, R., Fox, A.J., Bolton, F.E., Bootsma, H.J., Willems, R.J., Urwin, R. and Maiden, M.C., 2001. Multilocus sequence typing system for Campylobacter jejuni. *Journal of clinical microbiology*, 39(1), pp.14–23.

Dobzhansky, T., 1935. A Critique of the Species Concept in Biology. *Philosophy of Science*, 2(3), pp.344–355.

Doolittle, W.F. and Papke, R.T., 2006. Genomics and the bacterial species problem. *Genome biology*, 7(9), p.116.

Doolittle, W.F. and Zhaxybayeva, O., 2009. On the origin of prokaryotic species. *Genome Research*, 19(5), pp.744–756.

Eggleston, A.K., Mitchell, A.H. and West, S.C., 1997. In Vitro Reconstitution of the Late Steps of Genetic Recombination in E. coli. *Cell*, 89(4), pp.607–617.

Eisen, J.A., 2000. Horizontal gene transfer among microbial genomes: new insights from complete genome analysis. *Current opinion in genetics & development*, 10(6), pp.606–11.

Food, E. and Authority, S., 2019. The European Union summary report on antimicrobial resistance in zoonotic and indicator bacteria from humans, animals and food in 2017. *EFSA Journal*, 17(2).

Gaynor, E.C., Cawthraw, S., Manning, G., MacKichan, J.K., Falkow, S. and Newell, D.G., 2004. The Genome-Sequenced Variant of Campylobacter jejuni NCTC 11168 and the Original Clonal Clinical Isolate Differ Markedly in Colonization, Gene Expression, and Virulence-Associated Phenotypes. *Journal of Bacteriology*, 186(23), pp.8159–8159.

Gilbert, M.J., Kik, M., Miller, W.G., Duim, B. and Wagenaar, J.A., 2015. Campylobacter iguaniorum sp. nov., isolated from reptiles. *International Journal of Systematic and Evolutionary Microbiology*, 65(Pt_3), pp.975–982.

Gilbert, M.J., Miller, W.G., Leger, J.S., Chapman, M.H., Timmerman, A.J., Duim, B., Foster, G. and Wagenaar, J.A., 2017. Campylobacter pinnipediorum sp. nov., isolated from pinnipeds, comprising Campylobacter pinnipediorum subsp. pinnipediorum subsp. nov. and Campylobacter pinnipediorum subsp. caledonicus subsp. nov. *International Journal of Systematic and Evolutionary Microbiology*, 67(6), pp.1961–1968.

Griekspoor, P., Colles, F.M., McCarthy, N.D., Hansbro, P.M., Ashhurst-Smith, C., Olsen, B., Hasselquist, D., Maiden, M.C.J. and Waldenström, J., 2013. Marked host specificity and lack of phylogeographic population structure of Campylobacter jejuni in wild birds. *Molecular Ecology*, 22(5), pp.1463–1472.

Griekspoor, P., Olsen, B. and Waldenström, J., 2009. Campylobacter jejuni in penguins, Antarctica. *Emerging infectious diseases*, 15(5), pp.847–8.

Gupta, R.S., Lo, B. and Son, J., 2018. Phylogenomics and Comparative Genomic Studies Robustly Support Division of the Genus Mycobacterium into an

Emended Genus Mycobacterium and Four Novel Genera. *Frontiers in microbiology*, 9(FEB), p.67.

Haddad, N., Tresse, O., Rivoal, K., Chevret, D., Nonglaton, Q., Burns, C.M., Prévost, H. and Cappelier, J.M., 2012. Polynucleotide phosphorylase has an impact on cell biology of Campylobacter jejuni. *Frontiers in Cellular and Infection Microbiology*, 2(March), p.30.

Hanage, W.P., 2013. Fuzzy species revisited. *BMC Biology*, 11(1), p.41.

Hanage, W.P., Fraser, C. and Spratt, B.G., 2005. Fuzzy species among recombinogenic bacteria. *BMC biology*, 3, p.6.

Helgason, E., Caugant, D.A., Olsen, I. and Kolstø, A.-B., 2000. Genetic Structure of Population of Bacillus cereus and B. thuringiensis Isolates Associated with Periodontitis and Other Human Infections. *Journal of Clinical Microbiology*, 38(4), pp.1615–1622.

Hernández, J. and González-Acuña, D., 2016. Anthropogenic antibiotic resistance genes mobilization to the polar regions. *Infection Ecology & Epidemiology*, 6(1), p.32112.

Hernández, J., Stedt, J., Bonnedahl, J., Molin, Y., Drobni, M., Calisto-Ulloa, N., Gomez-Fuentes, C., Astorga-España, M.S., González-Acuña, D., Waldenström, J., Blomqvist, M. and Olsen, B., 2012. Human-Associated Extended-Spectrum β-Lactamase in the Antarctic. *Applied and Environmental Microbiology*, 78(6), pp.2056–2058.

Holt, R.D., 2009. Bringing the Hutchinsonian niche into the 21st century: Ecological and evolutionary perspectives. *Proceedings of the National Academy of Sciences*, 106(Supplement_2), pp.19659–19665.

Hormeño, L., Palomo, G., Ugarte-Ruiz, M., Porrero, M.C., Borge, C., Vadillo, S., Píriz, S., Domínguez, L., Campos, M.J. and Quesada, A., 2016. Identification of the main quinolone resistance determinant in Campylobacter jejuni and Campylobacter coli by MAMA-DEG PCR. *Diagnostic Microbiology and Infectious Disease*, 84(3), pp.236–239.

Howlett, R.M., Hughes, B.M., Hitchcock, A. and Kelly, D.J., 2012. Hydrogenase activity in the foodborne pathogen Campylobacter jejuni depends upon a novel ABC-type nickel transporter (NikZYXWV) and is SlyD-independent.

Microbiology, 158(6), pp.1645–1655.

Hyytiäinen, H., Juntunen, P., Scott, T., Kytömäki, L., Venho, R., Laiho, A., Junttila, S., Gyenesei, A., Revez, J. and Hänninen, M.-L., 2013. Effect of ciprofloxacin exposure on DNA repair mechanisms in Campylobacter jejuni. *Microbiology*, 159(Pt_12), pp.2513–2523.

Iraola, G., Forster, S.C., Kumar, N., Lehours, P., Bekal, S., García-Peña, F.J., Paolicchi, F., Morsella, C., Hotzel, H., Hsueh, P.-R., Vidal, A., Lévesque, S., Yamazaki, W., Balzan, C., Vargas, A., Piccirillo, A., Chaban, B., Hill, J.E., Betancor, L., Collado, L., Truyers, I., Midwinter, A.C., Dagi, H.T., Mégraud, F., Calleros, L., Pérez, R., Naya, H. and Lawley, T.D., 2017. Distinct Campylobacter fetus lineages adapted as livestock pathogens and human pathobionts in the intestinal microbiota. *Nature Communications*, 8(1), p.1367.

Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T. and Aluru, S., 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature Communications*, 9(1), p.5114.

Jia, B., Raphenya, A.R., Alcock, B., Waglechner, N., Guo, P., Tsang, K.K., Lago, B.A., Dave, B.M., Pereira, S., Sharma, A.N., Doshi, S., Courtot, M., Lo, R., Williams, L.E., Frye, J.G., Elsayegh, T., Sardar, D., Westman, E.L., Pawlowski, A.C., Johnson, T.A., Brinkman, F.S.L., Wright, G.D. and McArthur, A.G., 2017. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Research*, 45(D1), pp.D566–D573.

Jiang, X. and Fares, M.A., 2017. Two SecA / SecY Systems with Distinct Roles in the Ecological Adaptations of Bacteria. *bioRxiV*, (October).

Katoh, K., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30(14), pp.3059–3066.

Kirk, K.F., Méric, G., Nielsen, H.L., Pascoe, B., Sheppard, S.K., Thorlacius-Ussing, O. and Nielsen, H., 2018. Molecular epidemiology and comparative genomics of Campylobacter concisus strains from saliva, faeces and gut mucosal biopsies in inflammatory bowel disease. *Scientific Reports*, 8(1), p.1902.

Lapierre, L., Arias, M.L. and Fernández, H., 2016. Antimicrobial Resistance in

Campylobacter spp. *Campylobacter spp. and Related Organisms in Poultry*. Cham: Springer International Publishing, pp.165–183.

Lawson, D.J., Hellenthal, G., Myers, S. and Falush, D., 2012. Inference of Population Structure using Dense Haplotype Data. *PLoS Genetics*, 8(1), p.e1002453.

Livermore, D.M., 2007. Introduction: the challenge of multiresistance. *International Journal of Antimicrobial Agents*, 29(SUPPL. 3), pp.S1–S7.

Logan, J.M.J., Burnens, A., Linton, D., Lawson, A.J. and Stanley, J., 2000. Campylobacter lanienae sp. nov., a new species isolated from workers in an abattoir. *International Journal of Systematic and Evolutionary Microbiology*, 50(2), pp.865–872.

Luo, N., Sahin, O., Lin, J., Michel, L.O. and Zhang, Q., 2003. In Vivo Selection of Campylobacter Isolates with High Levels of Fluoroquinolone Resistance Associated with gyrA Mutations and the Function of the CmeABC Efflux Pump. *Antimicrobial Agents and Chemotherapy*, 47(1), pp.390–394.

Mallet, J., Besansky, N. and Hahn, M.W., 2016. How reticulated are species? *BioEssays*, 38(2), pp.140–149.

Mayr, E., 1942. *Systematics and the Origin of Species*. New York: Columbia University Press.

McInerney, J.O., Whelan, F.J., Domingo-Sananes, M.R., McNally, A. and O'Connell, M.J., 2020. Pangenomes and Selection: The Public Goods Hypothesis. *The Pangenome*. Cham: Springer International Publishing, pp.151–167.

McLennan, M.K., Ringoir, D.D., Frirdich, E., Svensson, S.L., Wells, D.H., Jarrell, H., Szymanski, C.M. and Gaynor, E.C., 2008. Campylobacter jejuni Biofilms Up-Regulated in the Absence of the Stringent Response Utilize a Calcofluor White-Reactive Polysaccharide. *Journal of Bacteriology*, 190(3), pp.1097–1107.

Méric, G., Yahara, K., Mageiros, L., Pascoe, B., Maiden, M.C.J., Jolley, K.A. and Sheppard, S.K., 2014. A Reference Pan-Genome Approach to Comparative Bacterial Genomics: Identification of Novel Epidemiological Markers in Pathogenic Campylobacter. *PLoS ONE*, 9(3), p.e92798.

Mikheil, D.M., Shippy, D.C., Eakley, N.M., Okwumabua, O.E. and Fadl, A.A., 2012. Deletion of gene encoding methyltransferase (gidB) confers high-level

antimicrobial resistance in Salmonella. *The Journal of Antibiotics*, 65(4), pp.185–192.

Miller, W.G., Yee, E., Chapman, M.H., Smith, T.P.L., Bono, J.L., Huynh, S., Parker, C.T., Vandamme, P., Luong, K. and Korlach, J., 2014. Comparative genomics of the Campylobacter lari group. *Genome biology and evolution*, 6(12), pp.3252–66.

Mourkas, E., Florez-Cuadrado, D., Pascoe, B., Calland, J.K., Bayliss, S.C., Mageiros, L., Méric, G., Hitchings, M.D., Quesada, A., Porrero, C., Ugarte-Ruiz, M., Gutiérrez-Fernández, J., Domínguez, L. and Sheppard, S.K., 2019. Gene pool transmission of multidrug resistance among Campylobacter from livestock, sewage and human disease. *Environmental Microbiology*, 21(12), pp.4597–4613.

Mourkas, E., Taylor, A.J., Méric, G., Bayliss, S.C., Pascoe, B., Mageiros, L., Calland, J.K., Hitchings, M.D., Ridley, A., Vidal, A., Forbes, K.J., Strachan, N.J.C., Parker, C.T., Parkhill, J., Jolley, K.A., Cody, A.J., Maiden, M.C.J., Kelly, D.J. and Sheppard, S.K., 2020. Agricultural intensification and the evolution of host specialism in the enteric pathogen Campylobacter jejuni. *Proceedings of the National Academy of Sciences*, 117(20), pp.11018–11028.

Murray, S., Pascoe, B., Méric, G., Mageiros, L., Yahara, K., Hitchings, M.D., Friedmann, Y., Wilkinson, T.S., Gormley, F.J., Mack, D., Bray, J.E., Lamble, S., Bowden, R., Jolley, K.A., Maiden, M.C.J., Wendlandt, S., Schwarz, S., Corander, J., Fitzgerald, J.R. and Sheppard, S.K., 2017. Recombination-Mediated Host Adaptation by Avian Staphylococcus aureus. *Genome Biology and Evolution*, 9(4), pp.830–842.

Ochman, H., Lawrence, J.G. and Groisman, E.A., 2000. Lateral gene transfer and the nature of bacterial innovation. *Nature*, 405(6784), pp.299–304.

Olsen, B., Bergström, S., McCafferty, D., Sellin, M. and Wiström, J., 1996. Salmonella enteritidis in Antarctica: zoonosis in man or humanosis in penguins? *The Lancet*, 348(9037), pp.1319–1320.

Page, A.J., Taylor, B., Delaney, A.J., Soares, J., Seemann, T., Keane, J.A. and Harris, S.R., 2016. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microbial Genomics*, 2(4), p.e000056.

Park, C.J. and Andam, C.P., 2020. Distinct but Intertwined Evolutionary Histories of Multiple Salmonella enterica Subspecies. *mSystems*, 5(1), pp.1–14.

Pascoe, B., Méric, G., Yahara, K., Wimalarathna, H., Murray, S., Hitchings, M.D., Sproston, E.L., Carrillo, C.D., Taboada, E.N., Cooper, K.K., Huynh, S., Cody, A.J., Jolley, K.A., Maiden, M.C.J., McCarthy, N.D., Didelot, X., Parker, C.T. and Sheppard, S.K., 2017. Local genes for local bacteria: Evidence of allopatry in the genomes of transatlantic Campylobacter populations. *Molecular Ecology*, 26(17), pp.4497–4508.

Pascoe, B., Williams, L.K., Calland, J.K., Meric, G., Hitchings, M.D., Dyer, M., Ryder, J., Shaw, S., Lopes, B.S., Chintoan-Uta, C., Allan, E., Vidal, A., Fearnley, C., Everest, P., Pachebat, J.A., Cogan, T.A., Stevens, M.P., Humphrey, T.J., Wilkinson, T.S., Cody, A.J., Colles, F.M., Jolley, K.A., Maiden, M.C.J., Strachan, N., Pearson, B.M., Linton, D., Wren, B.W., Parkhill, J., Kelly, D.J., van Vliet, A.H.M., Forbes, K.J. and Sheppard, S.K., 2019. Domestication of Campylobacter jejuni NCTC 11168. *Microbial genomics*, 5(7).

Piccirillo, A., Niero, G., Calleros, L., Pérez, R., Naya, H. and Iraola, G., 2016. Campylobacter geochelonis sp. nov. isolated from the western Hermann's tortoise (Testudo hermanni hermanni). *International Journal of Systematic and Evolutionary Microbiology*, 66(9), pp.3468–3476.

Pittman, M.S., Elvers, K.T., Lee, L., Jones, M.A., Poole, R.K., Park, S.F. and Kelly, D.J., 2007. Growth of Campylobacter jejuni on nitrate and nitrite: electron transport to NapA and NrfA via NrfH and distinct roles for NrfA and the globin Cgb in protection against nitrosative stress. *Molecular Microbiology*, 63(2), pp.575–590.

Potter, L., Angove, H., Richardson, D. and Cole, J., 2001. Nitrate reduction in the periplasm of gram-negative bacteria. *Advances in Microbial Physiology*. pp.51–112.

Qin, S., Wang, Y., Zhang, Q., Chen, X., Shen, Z., Deng, F., Wu, C. and Shen, J., 2012. Identification of a Novel Genomic Island Conferring Resistance to Multiple Aminoglycoside Antibiotics in Campylobacter coli. *Antimicrobial Agents and Chemotherapy*, 56(10), pp.5332–5339.

Reddy, S. and Zishiri, O.T., 2018. Genetic characterisation of virulence genes associated with adherence, invasion and cytotoxicity in *Campylobacter* spp.

isolated from commercial chickens and human clinical cases. *Onderstepoort Journal of Veterinary Research*, 85(1), pp.1–9.

Richardson, E.J., Bacigalupe, R., Harrison, E.M., Weinert, L.A., Lycett, S., Vrieling, M., Robb, K., Hoskisson, P.A., Holden, M.T.G., Feil, E.J., Paterson, G.K., Tong, S.Y.C., Shittu, A., van Wamel, W., Aanensen, D.M., Parkhill, J., Peacock, S.J., Corander, J., Holmes, M. and Fitzgerald, J.R., 2018. Gene exchange drives the ecological success of a multi-host bacterial pathogen. *Nature Ecology & Evolution*, 2(9), pp.1468–1478.

Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 30(14), pp.2068–2069.

Shapiro, B.J., Leducq, J.-B. and Mallet, J., 2016. What Is Speciation? *PLOS Genetics*, 12(3), p.e1005860.

Sheinman, M., Arkhipova, K., Arndt, P.F., Dutilh, B.E., Hermsen, R. and Massip, F., 2020. Long identical sequences found in multiple bacterial genomes reveal frequent and widespread exchange of genetic material between distant species. *bioRxiV*.

Sheppard, S.K., Cheng, L., Méric, G., de Haan, C.P.A., Llarena, A.-K., Marttinen, P., Vidal, A., Ridley, A., Clifton-Hadley, F., Connor, T.R., Strachan, N.J.C., Forbes, K., Colles, F.M., Jolley, K.A., Bentley, S.D., Maiden, M.C.J., Hänninen, M.-L., Parkhill, J., Hanage, W.P. and Corander, J., 2014. Cryptic ecology among host generalist Campylobacter jejuni in domestic animals. *Molecular Ecology*, 23(10), pp.2442–2451.

Sheppard, S.K., Colles, F.M., McCarthy, N.D., Strachan, N.J.C., Ogden, I.D., Forbes, K.J., Dallas, J.F. and Maiden, M.C.J., 2011. Niche segregation and genetic structure of Campylobacter jejuni populations from wild and agricultural host species. *Molecular ecology*, 20(16), pp.3484–90.

Sheppard, S.K., Dallas, J.F., Wilson, D.J., Strachan, N.J.C., McCarthy, N.D., Jolley, K.A., Colles, F.M., Rotariu, O., Ogden, I.D., Forbes, K.J. and Maiden, M.C.J., 2010. Evolution of an Agriculture-Associated Disease Causing Campylobacter coli Clade: Evidence from National Surveillance Data in Scotland. *PLoS ONE*, 5(12), p.e15708.

Sheppard, S.K., Didelot, X., Jolley, K.A., Darling, A.E., Pascoe, B., Meric, G., Kelly,

D.J., Cody, A., Colles, F.M., Strachan, N.J.C., Ogden, I.D., Forbes, K., French, N.P., Carter, P., Miller, W.G., McCarthy, N.D., Owen, R., Litrup, E., Egholm, M., Affourtit, J.P., Bentley, S.D., Parkhill, J., Maiden, M.C.J. and Falush, D., 2013. Progressive genome-wide introgression in agricultural Campylobacter coli. *Molecular ecology*, 22(4), pp.1051–64.

Sheppard, S.K., Didelot, X., Meric, G., Torralbo, A., Jolley, K.A., Kelly, D.J., Bentley, S.D., Maiden, M.C.J., Parkhill, J. and Falush, D., 2013. Genome-wide association study identifies vitamin B5 biosynthesis as a host specificity factor in Campylobacter. *Proceedings of the National Academy of Sciences of the United States of America*, 110(29), pp.11923–7.

Sheppard, S.K., Guttman, D.S. and Fitzgerald, J.R., 2018. Population genomics of bacterial host adaptation. *Nature Reviews Genetics*, 19(9), pp.549–565.

Sheppard, S.K., Jolley, K.A. and Maiden, M.C.J., 2012. A Gene-By-Gene Approach to Bacterial Population Genomics: Whole Genome MLST of Campylobacter. *Genes*, 3(2), pp.261–277.

Sheppard, S.K., McCarthy, N.D., Falush, D. and Maiden, M.C.J., 2008. Convergence of Campylobacter Species: Implications for Bacterial Evolution. *Science*, 320(5873), pp.237–239.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), pp.1312–1313.

Szymanski, C.M., Logan, S.M., Linton, D. and Wren, B.W., 2003. Campylobacter – a tale of two protein glycosylation systems. *Trends in Microbiology*, 11(5), pp.233–238.

Teuber, M., 2001. Veterinary use and antibiotic resistance. *Current Opinion in Microbiology*, 4(5), pp.493–499.

Thomas, C.M. and Nielsen, K.M., 2005. Mechanisms of, and Barriers to, Horizontal Gene Transfer between Bacteria. *Nature Reviews Microbiology*, 3(9), pp.711–721.

Tran, F. and Boedicker, J.Q., 2017. Genetic cargo and bacterial species set the rate of vesicle-mediated horizontal gene transfer. *Scientific Reports*, 7(1), p.8813.

Tuchscherr, L., Medina, E., Hussain, M., Völker, W., Heitmann, V., Niemann, S., Holzinger, D., Roth, J., Proctor, R.A., Becker, K., Peters, G. and Löffler, B.,

2011. Staphylococcus aureus phenotype switching: an effective bacterial strategy to escape host immune response and establish a chronic infection. *EMBO Molecular Medicine*, 3(3), pp.129–141.

Van, T.T.H., Elshagmani, E., Gor, M.C., Scott, P.C. and Moore, R.J., 2016. Campylobacter hepaticus sp. nov., isolated from chickens with spotty liver disease. *International Journal of Systematic and Evolutionary Microbiology*, 66(11), pp.4518–4524.

Vos, M. and Eyre-Walker, A., 2017. Are pangenomes adaptive or not? *Nature microbiology*, 2(12), p.1576.

Weinert, L.A., Chaudhuri, R.R., Wang, J., Peters, S.E., Corander, J., Jombart, T., Baig, A., Howell, K.J., Vehkala, M., Välimäki, N., Harris, D., Chieu, T.T.B., Van Vinh Chau, N., Campbell, J., Schultsz, C., Parkhill, J., Bentley, S.D., Langford, P.R., Rycroft, A.N., Wren, B.W., Farrar, J., Baker, S., Hoa, N.T., Holden, M.T.G., Tucker, A.W. and Maskell, D.J., 2015. Genomic signatures of human and animal disease in the zoonotic pathogen Streptococcus suis. *Nature Communications*, 6(1), p.6740.

Werren, J.H., 2011. Selfish genetic elements, genetic conflict, and evolutionary innovation. *Proceedings of the National Academy of Sciences*, 108(Supplement_2), pp.10863–10870.

Wilson, D.J., Gabriel, E., Leatherbarrow, A.J.H., Cheesbrough, J., Gee, S., Bolton, E., Fox, A., Hart, C.A., Diggle, P.J. and Fearnhead, P., 2009. Rapid Evolution and the Importance of Recombination to the Gastroenteric Pathogen Campylobacter jejuni. *Molecular Biology and Evolution*, 26(2), pp.385–397.

Wösten, M.M.S.M., Dubbink, H.J. and van der Zeijst, B.A.., 1996. The aroA gene of Campylobacter jejuni. *Gene*, 181(1–2), pp.109–112.

Yahara, K., Didelot, X., Ansari, M.A., Sheppard, S.K. and Falush, D., 2014. Efficient Inference of Recombination Hot Regions in Bacterial Genomes. *Molecular Biology and Evolution*, 31(6), pp.1593–1605.

Yahara, K., Didelot, X., Jolley, K.A., Kobayashi, I., Maiden, M.C.J., Sheppard, S.K. and Falush, D., 2016. The Landscape of Realized Homologous Recombination in Pathogenic Bacteria. *Molecular Biology and Evolution*, 33(2), pp.456–471.

Yahara, K., Méric, G., Taylor, A.J., de Vries, S.P.W., Murray, S., Pascoe, B.,

Mageiros, L., Torralbo, A., Vidal, A., Ridley, A., Komukai, S., Wimalarathna, H., Cody, A.J., Colles, F.M., McCarthy, N., Harris, D., Bray, J.E., Jolley, K.A., Maiden, M.C.J., Bentley, S.D., Parkhill, J., Bayliss, C.D., Grant, A., Maskell, D., Didelot, X., Kelly, D.J. and Sheppard, S.K., 2017. Genome-wide association of functional traits linked with C ampylobacter jejuni survival from farm to fork. *Environmental Microbiology*, 19(1), pp.361–380.

Yahara, K., Nakayama, S., Shimuta, K., Lee, K., Morita, M., Kawahata, T., Kuroki, T., Watanabe, Y., Ohya, H., Yasuda, M., Deguchi, T., Didelot, X. and Ohnishi, M., 2018. Genomic surveillance of Neisseria gonorrhoeae to investigate the distribution and evolution of antimicrobial-resistance determinants and lineages. *Microbial Genomics*, 4(9), pp.1–3.

Yao, H., Shen, Z., Wang, Y., Deng, F., Liu, D., Naren, G., Dai, L., Su, C., Wang, B., Wang, S., Wu, C., Yu, E.W., Zhang, Q. and Shen, J., 2016. Emergence of a Potent Multidrug Efflux Pump Variant That Enhances Campylobacter Resistance to Multiple Antibiotics. *mBio*, 7(5), pp.1–11.

Young, J.P.W., 2016. Bacteria Are Smartphones and Mobile Genes Are Apps. *Trends in microbiology*, 24(12), pp.931–932.

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., Aarestrup, F.M. and Larsen, M.V., 2012. Identification of acquired antimicrobial resistance genes. *Journal of Antimicrobial Chemotherapy*, 67(11), pp.2640–2644.

Zhu, P., van der Ende, A., Falush, D., Brieske, N., Morelli, G., Linz, B., Popovic, T., Schuurman, I.G.A., Adegbola, R.A., Zurth, K., Gagneux, S., Platonov, A.E., Riou, J.-Y., Caugant, D.A., Nicolas, P. and Achtman, M., 2001. Fit genotypes and escape variants of subgroup III Neisseria meningitidis during three pandemics of epidemic meningitis. *Proceedings of the National Academy of Sciences*, 98(9), pp.5234–5239.

Data accessibility

Contiguous assemblies of all genome sequences compared are available at the public data repository Figshare (doi.org/10.6084/m9.figshare.12741296). Individual accession numbers can be found in Table S2.1 at the Appendix.

Chapter 3

Agricultural intensification and the evolution of host specialism in the enteric pathogen *Campylobacter jejuni*

The work presented in this chapter is published as a peer-reviewed publication at:

Mourkas, E., Taylor, J.T., Méric, G., Bayliss, C.S., Pascoe, B., Mageiros, L., Calland, K.C., Hitchings, M.D., Ridley, A., Vidal, A., Forbes, J.K., Strachan, J.C.N., Parker, T.C., Parkhill, J., Jolley, A.K., Cody, J.A., Maiden, C.J.M., Kelly, J.D., Sheppard, S.K. 2020. "Agricultural intensification and the evolution of host specialism in the enteric pathogen *Campylobacter jejuni*." *PNAS*. doi: 10.1073/pnas.1917168117.

Commentary text

The work in this chapter builds on work from the previous chapter by delving into the evolution of host specific lineages within the most commonly isolated species in the *Campylobacter* genus. This chapter provides a comprehensive analysis of a dataset of 1,200 *C. jejuni* isolates from several sources, countries and isolation years. Comparative genomics and phylogenetic analyses identified a dominant cattle-associated cluster (ST-61 clonal complex) that emerged from a background of ecological generalists around the time of intensification of modern beef production. This host-specific genome evolution was mirrored among accessory genes where reductive evolution and the loss of a particular genomic island was an important factor in cattle specialization. The work in this chapter provides evidence of adaptation, following multiple host transitions to cattle and has important implications for the emergence and dissemination of novel pathogenic clones in modern agricultural systems. The statement of authorship for this chapter can be found in the Appendix supplementary form SF2.

Abstract

Modern agriculture has dramatically changed the distribution of animal species on earth. Changes to host ecology have a major impact on the microbiota, potentially increasing the risk of zoonotic pathogens being transmitted to humans, but the impact of intensive livestock production on host-associated bacteria has rarely been studied. Work in this chapter uses large isolate collections and comparative genomics techniques, linked to phenotype studies, to understand the timescale and genomic adaptations associated with the proliferation of the most common food-born bacterial pathogen (*C. jejuni*) in the most prolific agricultural mammal (cattle). Findings in this chapter reveal the emergence of cattle specialist *C. jejuni* lineages from a background of host generalist strains that coincided with the dramatic rise in cattle numbers in the 20th century. Cattle adaptation was associated with horizontal gene transfer and significant gene gain and loss. This may be related to differences in host diet, anatomy and physiology, leading to the proliferation of globally disseminated cattle specialists of major public health importance. This work highlights how genomic plasticity can allow important zoonotic pathogens to exploit altered niches in the face of anthropogenic change and provides information for mitigating some of the risks posed by modern agricultural systems.

Introduction

In the relatively short history since the second agricultural revolution (17th to 19th centuries), intensive farming practices have dramatically changed the distribution of plants and animals on earth. Today, livestock account for 60% of all mammal biomass, with cattle being the most abundant (Bar-On et al., 2018), far surpassing the biomass of all wild mammals combined (4%). This radical anthropogenic ecological change carries significant risk, and in recent decades, escalating livestock numbers and global trade have been linked with the emergence of zoonotic diseases that pose a significant threat to both animal and human health (Jones et al., 2013; Daszak, 2000; Cutler et al., 2010).

The most abundant human bacterial pathogen found in cattle is *Campylobacter* (Manyi-Loh et al., 2016). These organisms inhabit the gastrointestinal tract of many warm-blooded animals and are present in the faeces of around 20% of cows at a concentration of approximately $3x10^4$ CFU/g (Ogden et al., 2009). The most frequently isolated species, *C. jejuni* and *C. coli,* are the leading cause of bacterial

gastroenteritis in high income countries (Eurosurveillance editorial team, 2015; Friedman, C.R., Neiman, J., Wegener, H.C., 2000). These organisms typically infect humans via consumption of contaminated meat and poultry (Sheppard et al., 2009), leading to widespread morbidity, and occasionally mortality in vulnerable groups (Eurosurveillance editorial team, 2015; Friedman, C.R., Neiman, J., Wegener, H.C., 2000). There are an estimated 1.5 billion cattle on earth, each of which produces around 30 kg of manure daily (Bar-On et al., 2018). Therefore, around $3x10^{17}$ (300 quadrillion) *Campylobacter* are excreted by cattle into the environment every day. The sheer magnitude of shedding is clearly important in terms of direct environmental contamination (Manyi-Loh et al., 2016; Mourkas et al., 2019) and potential spillover into the human food chain. However, while much research has focussed on the epidemiology of transient human infection (Kaakoush et al., 2015), little is known about how intensive farming may establish new transmission cycles among reservoir hosts and influence the evolution of livestock adapted strains.

Spillover between reservoir host species is exemplified in *Campylobacter* by the existence of strains that are regularly isolated from multiple host species (ecological generalists), indicating recent host transitions (Woodcock et al., 2017). While frequent host transition and genome plasticity in *Campylobacter* can favour ecological generalism (Woodcock et al., 2017), expansion into a new host species is typically thought to be accompanied by evolutionary specialization and gradual divergence from the ancestral population. This process may be expedited in intensive farming where frequent animal contact, high animal numbers, and low genetic diversity provide opportunities for pathogens to evolve (Jones et al., 2013; McMichael, 2004; Daszak, 2000; Cutler et al., 2010), potentially promoting the emergence and proliferation of specialist lineages that exploit the niche more effectively. Furthermore, as host animal numbers and colonization increase so does the population size of the associated bacteria, enhancing the efficiency by which natural selection favours livestock adapted strains. In the context of increasingly intensive farming practices and extensive global trade networks, the shifting nature of livestock pathogens presents a major public health threat. For example, around half a million confirmed annual cases of human campylobacteriosis in Europe are caused by the two most common cattle associated *C. jejuni* genotypes (sequence types, ST-61 and ST-42) (Sheppard et al., 2014; Kwan et al., 2008; Sheppard et al., 2009), with the actual infection rate potentially much higher (Thépault et al., 2017).

This chapter carries out a population-genomic analysis of 1,198 *C. jejuni* isolates chosen to represent known diversity among reservoir host species in order to describe the nature and time-scale of adaptation to cattle. The work in this chapter aims to answer the following questions: (1) How do cattle-associated *C. jejuni* lineages emerge? (2) What is the timescale of their emergence? (3) What are the genomic elements associated with adaptation into cattle and (4) Can genotype-phenotype be linked mechanistically? The data reveal a dynamic pattern of genome evolution coinciding with the intensification of livestock farming and explain key adaptive processes of gene loss and gain that allow strains to exploit the metabolic and nutritional environment of the cow gut, and cope with different immune pressures. These evolutionary processes explain the emergence of one of the most common food-borne bacterial pathogens.

Results

Cattle specialist *C. jejuni* emerged from host generalist ancestors.

C. jejuni strains representing the breadth of the known genotype and host-source diversity (Appendix, Table S3.1) included 1,198 isolates from 18 different sources and belonging to 36 clonal complexes based upon sharing of 4 or more alleles at 7 housekeeping gene loci defined by multi-locus sequence typing (Dingle et al., 2001). A maximum-likelihood phylogenetic tree based on the gene-by-gene sequence alignment of 1,418 genes present in >90% of isolates, revealed genome sequence clusters that broadly corresponded to multilocus sequence typing (MLST) clonal complexes, consistent with previous studies (Figure 3.1) (Sheppard et al., 2014; Sheppard et al., 2013; Thépault et al., 2017). The phylogeny indicated broad diversity of isolates with evidence of host specialist (chicken or cattle) and host generalist lineages (e.g. ST-21 and ST-45 complexes) (Sheppard et al., 2014; Sheppard et al., 2013) (Figure 3.1). Consistent with host switching, isolates from cattle were common in generalist clades but 36% were found in discrete cattle associated clusters (ST-61 and ST-42 complexes). Strikingly, most ST-61 complex isolates (98%) formed a tight cluster at the tip of a longer branch emanating from within the ST-21 complex (Figure 3.1, Figure 3.2), potentially indicative of rapid differentiation from the ancestral population. To support evidence of the rapid expansion of this lineage the Tajima's D was calculated for the alignment of ST-61 complex isolate genomes (Frisch et al., 2018; Graña-Miraglia et al., 2017). This

generated a negative value (-1.569) implying an abundance of low frequency polymorphisms consistent with rapid clonal expansion of the population.

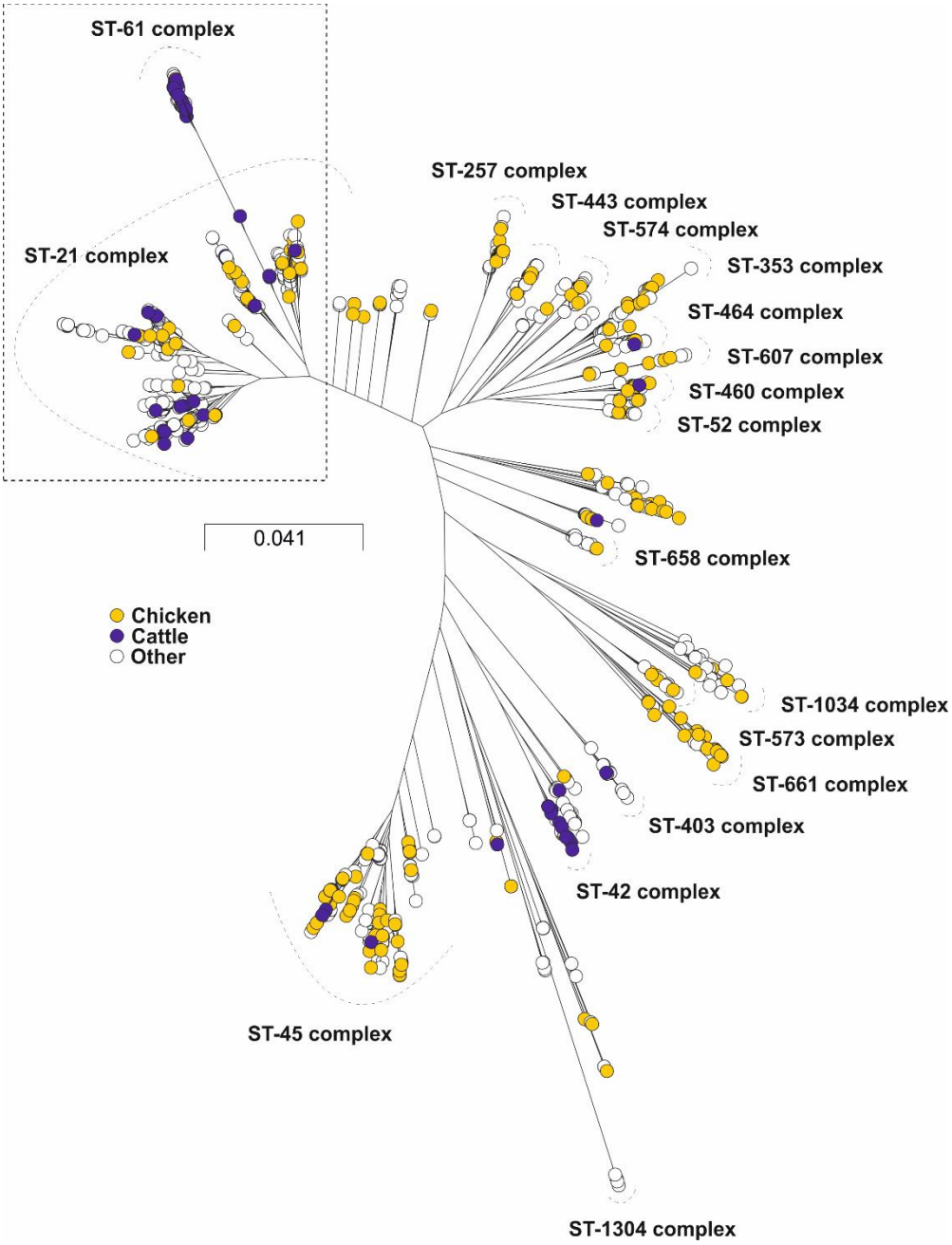


Figure 3.1. Cattle specialist *C. jejuni* emerged from a background of host generalists. (A) 1,198 *C. jejuni* isolates from chicken (yellow), ruminants (blue) and other sources (white) are shown on a phylogenetic tree reconstructed using an approximation of the maximum-likelihood algorithm (ML) implemented in RAxML, with the major MLST clonal complexes indicated next to the associated genome sequence cluster. The cattle specialist ST-61 clonal complex can be seen emerging from the generalist ST-21 complex (dashed box). The scale bar indicates the estimated number of substitutions per site.

Demographic reconstruction analysis using BEAST2 (Bouckaert et al., 2019) showed an increase in effective population size consistent with a scenario of rapid expansion of this lineage (Figure 3.3). Comparable analysis performed on ST-42

complex isolates showed no evidence of rapid lineage expansion based on Tajima's D (0.2) or demographic reconstruction (Figure 3.3).

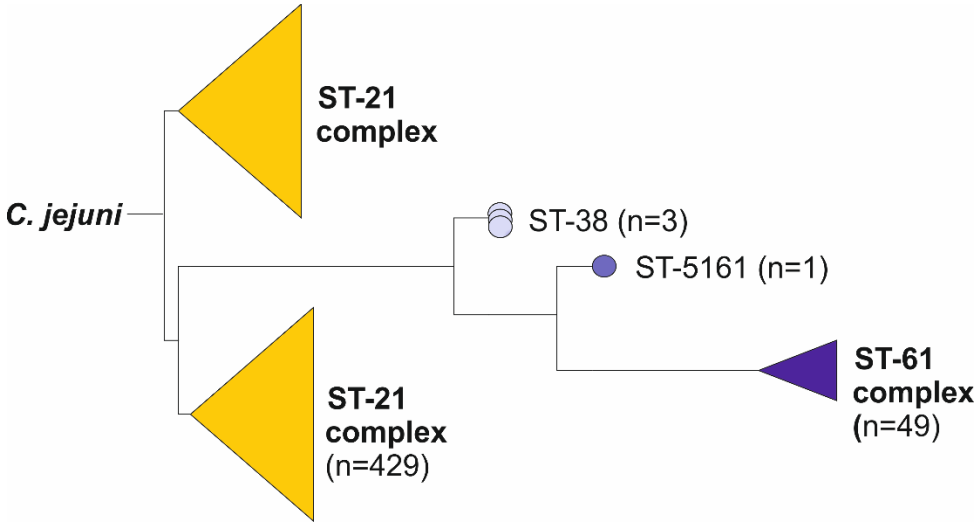


Figure 3.2. ML tree of the branch differentiating ST-61 complex (blue) from ST-21 complex (yellow) isolates highlighting the existence of intermediate isolates (different shades of blue) indicative of step-wise cattle specialization.

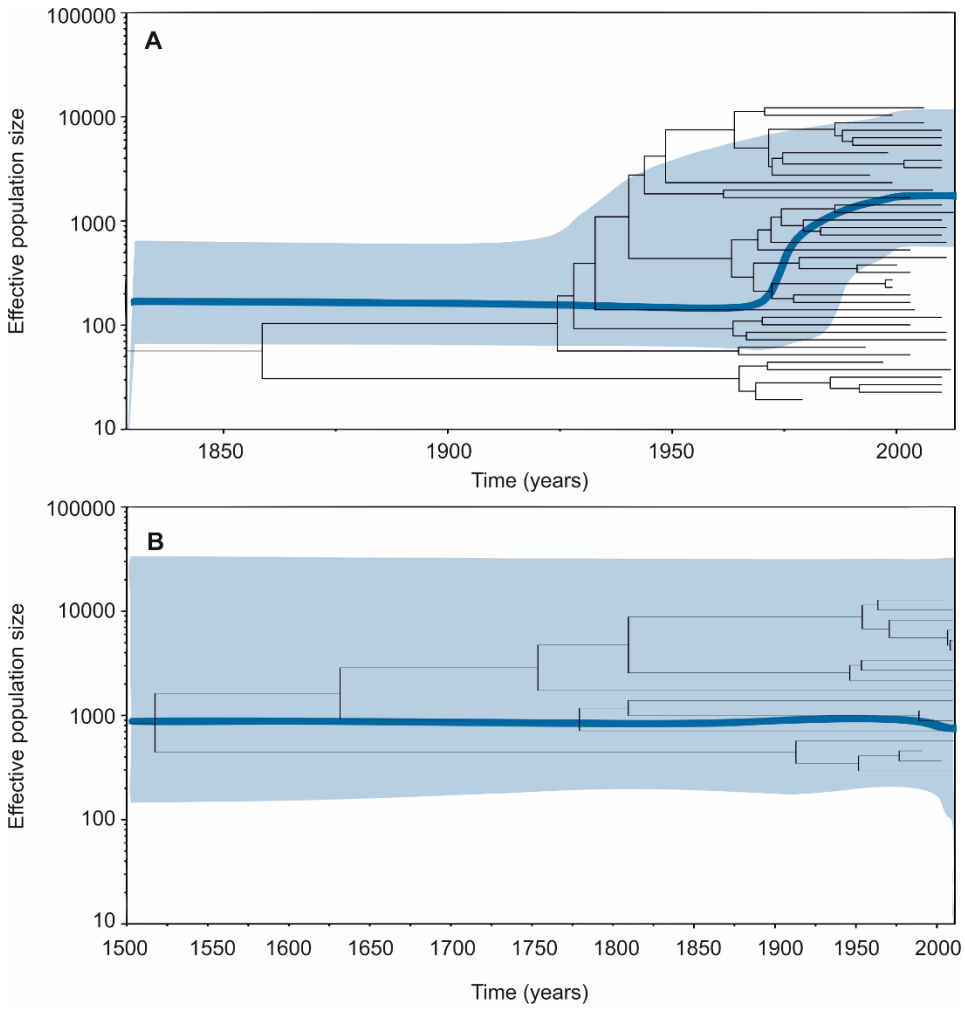


Figure 3.3. Effective population size of cattle specialists ST-61 and ST-42 phylogenies over time. The plot displays the effective population size and the generation length in years (Ne) of the

ST-61 (A) and ST-42 (B) lineages in relation to their phylogenies. The solid blue line shows the median effective population size over time, while the shaded blue coloured area indicates the 95% Highest Posterior Density. The effective population size (Ne) is shown on the *y* axis and the timeframe in years between the most recent sampling date and the emergence of each lineage on the *x* axis. The grey line indicates the line to the most recent common ancestor shared with the outgroup (MRCA).

Four isolates from cattle represented intermediates between the ST-21 and ST-61 complex sequence clusters (Figure 3.2). Comparison of the allelic variation in 43 genes that showed >90% allele segregation by complex (Appendix, Figure S3.1), revealed a gradual increase in 'ST-61-like' alleles from the first intermediate isolates (ST-38, 11/43) to the second intermediate (ST-5161, 27/43). The frequency distribution of average nucleotide identities for 1,208 core and 290 soft core ST-61 and ST-21 complex genes in the intermediate isolates revealed gradual differentiation from the ancestral ST-21 population, with a median core genome nucleotide identity decreasing from 99.47 (ST-38 isolates) to 99.25 (ST-5161 isolate) (Figure 3.2, Figure 3.4). These findings are consistent with the step-wise emergence of a cattle associated lineage.

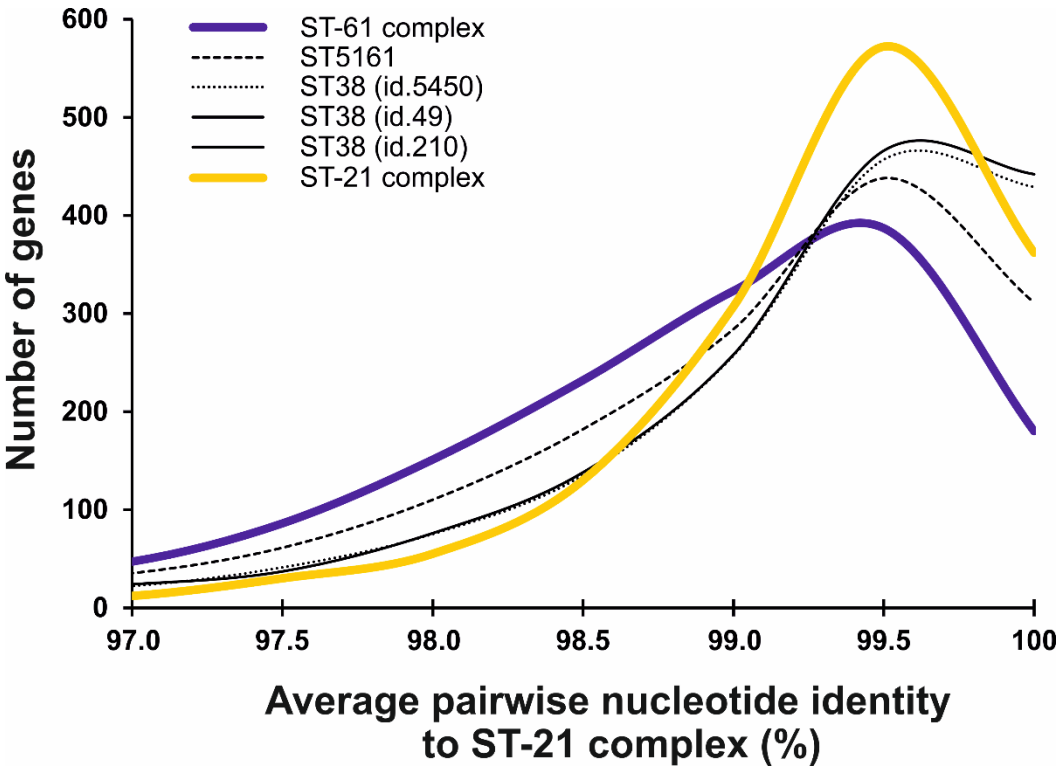


Figure 3.4. Average nucleotide identity for pairwise comparisons of 1,208 core and 290 soft core genes for 46 genomes of ST-21 complex (yellow), 49 genomes of ST-61 complex (blue) and 4 intermediate isolates from ST-38 and ST-5161 (black line, black dotted and black dotted respectively).

Intensive cattle farming coincides with the emergence of host specialist *C. jejuni*

To estimate the age of the emergence of the most common cattle specialist lineage, 41 ST-61 complex isolates with known isolation date ranging from 1979 to 2013, were used. A 1,490,602 bp core genome alignment of the 41 ST-61 complex isolates (90.81% of the NCTC11168 *C. jejuni* reference genome) was ordered on NCTC11168. Inferred recombinant regions were stringently masked to recover the population clonal frame and improve the timed-measured approximation of the phylogeny. The temporal signal of the phylogeny was calculated using TempEst v1.5.1 (Rambaut et al., 2016) (correlation coefficient ranged from 0.411 to 0.614) for ST-61 and ST-42 complexes and was consistent with other estimates for *C. jejuni* (Llarena et al., 2016). After removal of the recombined regions only variable sites remained. Ancestral dates were estimated for internal nodes using the tip dates of the ST-61 phylogeny and using a relaxed log normal clock model in BEAST2 (Bouckaert et al., 2019). The dated phylogeny approximated the emergence of the ST-61 complex to have occurred in 1859 (95% Highest Posterior Density: 1692 – 1943) (Figure 3.5).

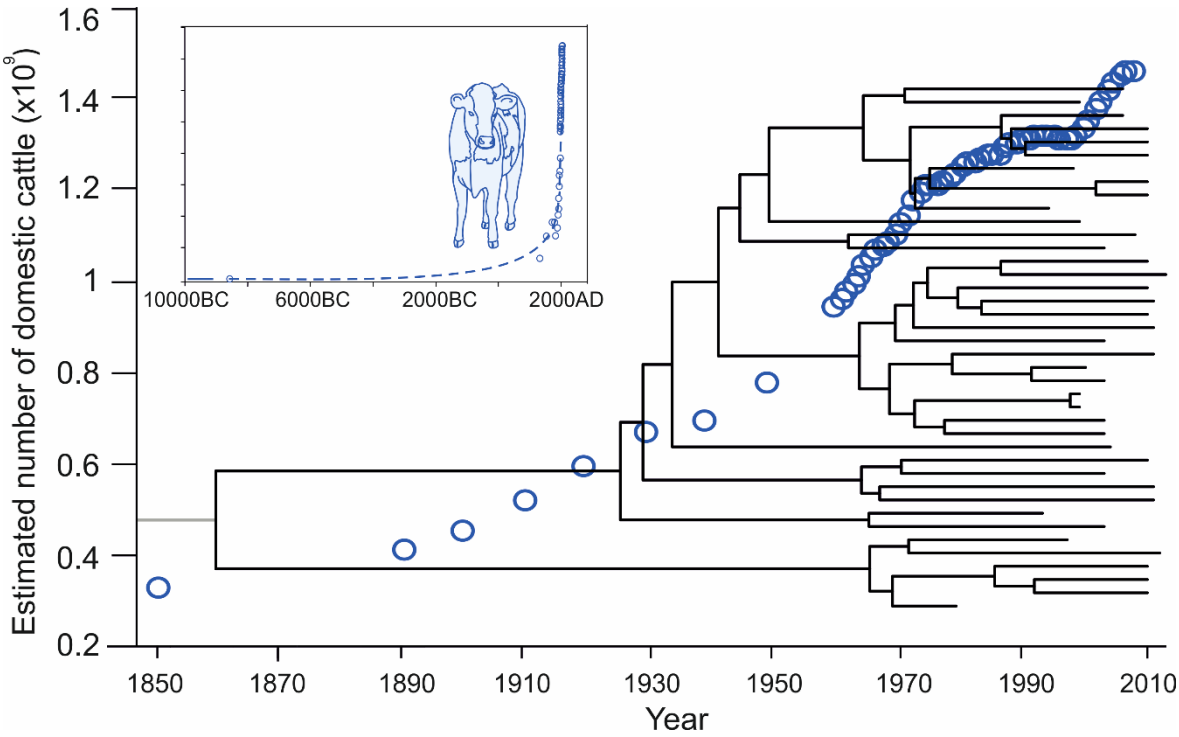


Figure 3.5. The emergence of cattle specialist *C. jejuni* coincides with modern intensive livestock farming. Graph of the estimated number of domestic cattle (blue circles) on earth from 1850 to 2010 (inset box from 10000BC to 2000AD) based upon data from (Bollongino et al., 2012; Allen, 2005; Mitchell, 1998; FAO, 2004). Time-scaled phylogeny of 41 *C. jejuni* isolates showing

dating estimates of the ST-61 complex emergence with ST-5161 as an outgroup, indicated as a gray line. The MRCA of ST-61 and ST-5161 was estimated.

Much of the ST-61 sublineage diversification was more recent (1925 to 1965) and there was evidence that population expansion coincided with the intensification of cattle farming throughout the 20th century. The same analysis was replicated for ST-42 complex where the time to most recent common ancestor (tMRCA) was dated in 1514 (95% Highest Posterior Density: 6647 BC – 1943) (Figure 3.3). Sublineages within the ST-42 complex have emerged from 1629 to 1954 (Figure 3.3).

Accessory gene gain is associated with cattle specialization

Both gene gain and gene loss can contribute to pathogen evolution and host tropism (Kettler et al., 2007; Koskiniemi et al., 2012) and this has been observed in *Campylobacter* (Morley et al., 2015; Sheppard et al., 2013). Comparative genomic analysis revealed 1,225 core genes present in all 1,198 isolates and 2,629 accessory genes that were differentially present. The prevalence of accessory genes varied between lineages (Figure 3.6) and was correlated with isolate host source. For example, a total of 35 genes were present in >90% of all ST-61 complex isolates (n=44/49) but absent in >90% of all ST-21 complex isolates (n=7/113) (Figure 3.7, Appendix, Table S3.2). This analysis was also performed to compare ST-42 and ST-21 complex isolates (Appendix, Table S3.3). While this does not prove that these genes confer an advantage when colonizing cattle, they can be considered to represent candidate adaptive genes. Most cattle associated accessory genes were annotated as hypothetical proteins of unknown function due to a lack of homology with clearly annotated and well characterized genes from an appropriate laboratory reference strain (Pascoe et al., 2019). Where the function could be inferred, genes were related to motility associated factors (*maf* family genes), involved in flagellar biosynthesis (*cj1340* and *cj1341c*) (Guerry et al., 2006; Novik et al., 2010), and capsule biosynthesis (*id86_0802* and *id86_0803*) (Table 3.1, Appendix, Table S3.3). Flagellar motility and surface structures are known to impact colonisation of animals *in vivo* (Wassenaar et al., 1993; Nachamkin et al., 1993; Maue et al., 2013), providing some support for a possible adaptive role for these genes in cattle colonization. However, *in vivo* assays would be necessary to confirm this.

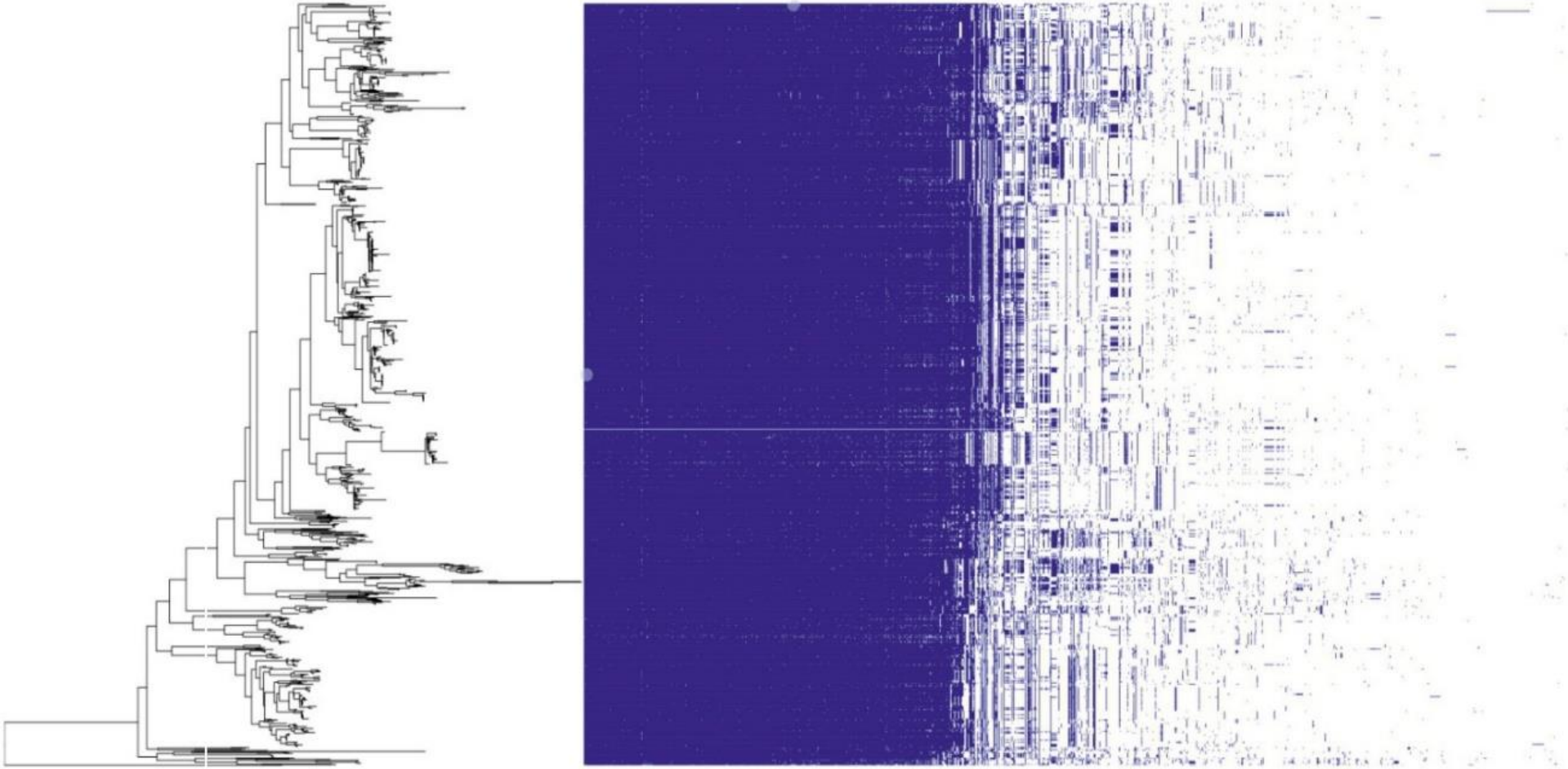


Figure 3.6. Phylogenetic tree structure of all 1,198 *C. jejuni* isolates and gene presence /absence matrix per isolate.

Table 3.1. Candidate adaptive genes involved in ST-61 complex emergence from ST-21 clonal complex.

Name	Alias	Transcri-ptional unit number [a]	Genome position on *C. jejuni* NCTC11168 reference genome	Predicted function [b]	COG family [b]	COG family description [b]	Prevalence in ST-21 complex (n=113)	Prevalence in ST-61 complex (n=49)	Prevalence in ST-42 complex (n=22)
Accessory genes highly prevalent in ST-61 and ST-42 complexes but missing in ST-21 complex									
CJM1_RS06530	-	-	-	Hypothetical protein	-	-	2/113 (1.77%)	44/49 (89.80%)	18/22 (81.82%)
CJM1_RS06535	-	-	-	Plasmid stabilization system protein	R	General function prediction only	2/113 (1.77%)	44/49 (89.80%)	18/22 (81.82%)
id1956_0137	*cj1340 -like*	-	-	Motility accessory factor (homolog)	S	Function unknown	0/113 (0%)	45/49 (91.84%)	21/22 (95.45%)
id3193_0984	*cj1341 -like*	-	-	Motility accessory factor (homolog)	S	Function unknown	0/113 (0%)	44/49 (89.80%)	20/22 (90.91%)
id380_0014	-	-	-	Hypothetical protein	-	-	4/113 (3.54%)	48/49 (97.96%)	22/22 (100%)
id5316_0084	-	-	-	Putative McrB subunit of the McrBC restriction endonuclease system	-	-	0/113 (0%)	48/49 (97.96%)	15/22 (68.18%)
Accessory genes highly prevalent in ST-21 complex but missing in ST-61 and ST-42 complexes									
cj1319	-	495	1248624.. 1249595	Putative nucleotide sugar dehydratase	M	Cell wall/membrane biogenesis	113/113 (100%)	1/49 (2.04%)	0/22 (0%)
cj1320	-	495	1249588.. 1250742	Putative DegT family aminotransferase	M	Cell wall/membrane biogenesis	113/113 (100%)	2/49 (4.08%)	0/22 (0%)
cj1324	-	498	1252278.. 1253399	Hypothetical protein	D	Cell cycle control, mitosis and meiosis	112/113 (99.1%)	1/49 (2.04%)	0/22 (0%)
cj1325	-	498	1253417.. 1254092	Putative methyltransferase	J	Translation	112/113 (99.1%)	1/49 (2.04%)	0/22 (0%)
cj1327	*neuB2*	498	1254132.. 1255136	N-acetylneuraminic acid synthetase	M	Cell wall/membrane biogenesis	113/113 (100%)	2/49 (4.08%)	0/22 (0%)
cj1328	*neuC2*	498	1255129.. 1256283	Putative UDP-N-acetylglucosamine 2-epimerase	M	Cell wall/membrane biogenesis	113/113 (100%)	1/49 (2.04%)	0/22 (0%)

cj1329	-	498	1256292..1257317	Putative sugar-phosphate nucleotide transferase	J	Translation	113/113 (100%)	1/49 (2.04%)	0/22 (0%)
cj1330	-	498	1257314..1258219	Hypothetical protein	R	General function prediction only	113/113 (100%)	1/49 (2.04%)	0/22 (0%)
cj1331	*ptmB*	498	1258212..1258919	Acylneuraminate cytidylyltransferase	M	Cell wall/membrane biogenesis	113/113 (100%)	2/49 (4.08%)	0/22 (0%)
cj1332	*ptmA*	498	1258919..1259689	Flagellin modification protein A	I	Lipid transport and metabolism	113/113 (100%)	2/49 (4.08%)	0/22 (0%)
Alleles of core genes harbouring recombination events highly prevalent in ST-61 and ST-42 complexes but missing in ST-21 complex									
cj0641	*pnk*	243	602326..603186	inorganic polyphosphate/ATP-NAD kinase	G	Carbohydrate transport and metabolism	4/113 (3.54%)	39/49 (79.59%)	20/22 (90.91%)
cj0665c	*argG*	248	621392..622612	argininosuccinate synthase	E	Amino acid transport and metabolism	4/113 (3.54%)	48/49 (97.96%)	21/22 (95.45%)
cj0709	*ffh*	269	665788..667125	signal recognition particle protein	U	Intracellular trafficking and secretion	4/113 (3.54%)	49/49 (100%)	20/22 (90.91%)
cj1044c	*thiH*	395	977626..978771	thiamine biosynthesis protein ThiH	H	Coenzyme transport and metabolism	0/113 (0%)	49/49 (100%)	22/22 (100%)
cj1082c	*thiD*	405	1013374..1014186	phosphomethylpyrimidi-ne kinase	H	Coenzyme transport and metabolism	6/113 (5.31%)	39/49 (79.59%)	22/22 (100%)
cj1349c	-	500	1280992..1282299	putative fibronectin/fibrinogen-binding protein	K	Transcription	0/113 (0%)	47/49 (95.92%)	17/22 (77.27%)
cj1350	*mobA*	501	1282361..1282936	putative molybdopterin-guanine dinucleotide biosynthesis protein	H	Coenzyme transport and metabolism	3/113 (2.65%)	42/49 (85.71%)	19/22 (86.36%)
cj1352	*ceuB*	502	1284008..1284976	enterochelin uptake permease	P	Inorganic ion transport and metabolism	3/113 (2.65%)	48/49 (97.96%)	22/22 (100%)
cj1633	-	601	1558654..1559637	putative ATP-binding protein	J	Translation	0/113 (0%)	44/49 (89.80%)	21/22 (95.45%)

ST-38 and ST-5161 intermediate isolates were not included in the prevalence calculations. One-letter abbreviations are linked to the functional categories in column 7 as defined by the Clusters of Orthologous Groups (COG) classification (Tatusov, 2000).
a. As predicted by the Prokaryotic Operon DataBase (ProOpDB) (Taboada et al., 2012). Consecutive numbers reflect consecutive transcriptional units on the *C. jejuni* NCTC11168 reference genome. b. As predicted by the WebMGA server (Wu et al., 2011).

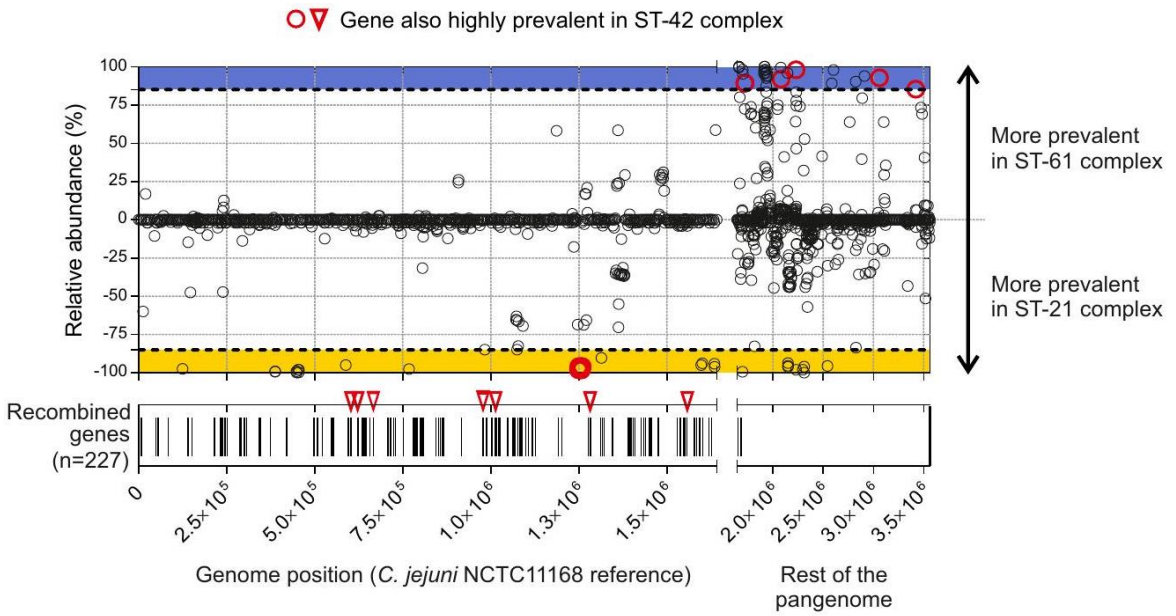


Figure 3.7. The cattle specialist lineage genomes show evidence of host adaptation. Gene presence/absence (black circles) and recombination regions (black bars) gained and lost on the tree branch differentiating ST-61 complex cattle specialist *C. jejuni* inferred using ClonalFrameML and mapped to the NCTC11168 reference genome. Accessory genes not mapped to NCTC11168 are also highlighted on the left (black circles – upper panel). The relative abundance of genes was calculated as presence in ST-61 complex isolates minus presence in ST-21 complex isolates. The frequency of all genes/alleles (black circles/bars) is shown for ST-61 and ST-21 complexes (upper and bottom panel). The homoplasic genes (red circles), including the glycosylation gene block, and recombinant alleles (red triangles) present or absent in ST-42 complex are shown (upper and bottom panel).

Homologous recombination and cattle specialization

A broadly accepted evolutionary principle in microbiology is that host-associated bacterial lineages with reduced genomes typically derive from ancestors with larger genomes (Ochman, 2005), because of mutational bias towards deleting superfluous or redundant genes (Mira et al., 2001). There is also evidence that genome reduction is enhanced when there is isolation from the ancestral population so recombination cannot mitigate against the accumulation of deleterious mutations (Sheppard et al., 2018). There was no evidence of genome reduction among ST-61 complex isolates sampled from 1979 to 2013 (Figure 3.8) or in the complex as a whole (1527-1835 genes) compared to ST-21 (1527-1885 genes) but this could be because of the relatively recent emergence of this cattle associated lineage. To investigate the extent of gene pool isolation, the extent of allele sharing (1,308 core genes) for 220 isolate genomes from clonal complexes that have been sampled from both chickens and cattle, was quantified. In 6 out of 9 clonal complexes the

average number of ST-61 alleles in the core genome was higher in isolates from cattle compared to those from chickens (Figure 3.9). This is consistent with homologous recombination among lineages that cohabit in cattle.

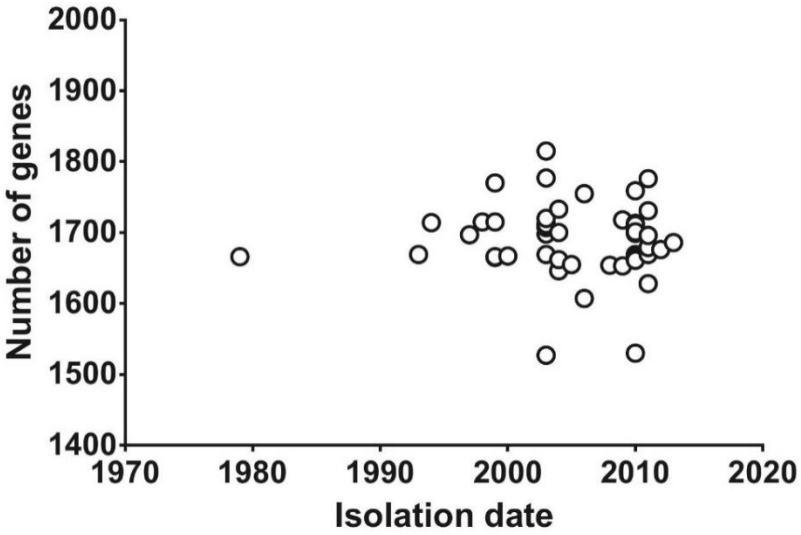


Figure 3.8. Number of genes of all ST-61 isolates. The total number of genes for every isolate is illustrated in the *y* axis, with isolation date on the *x* axis.

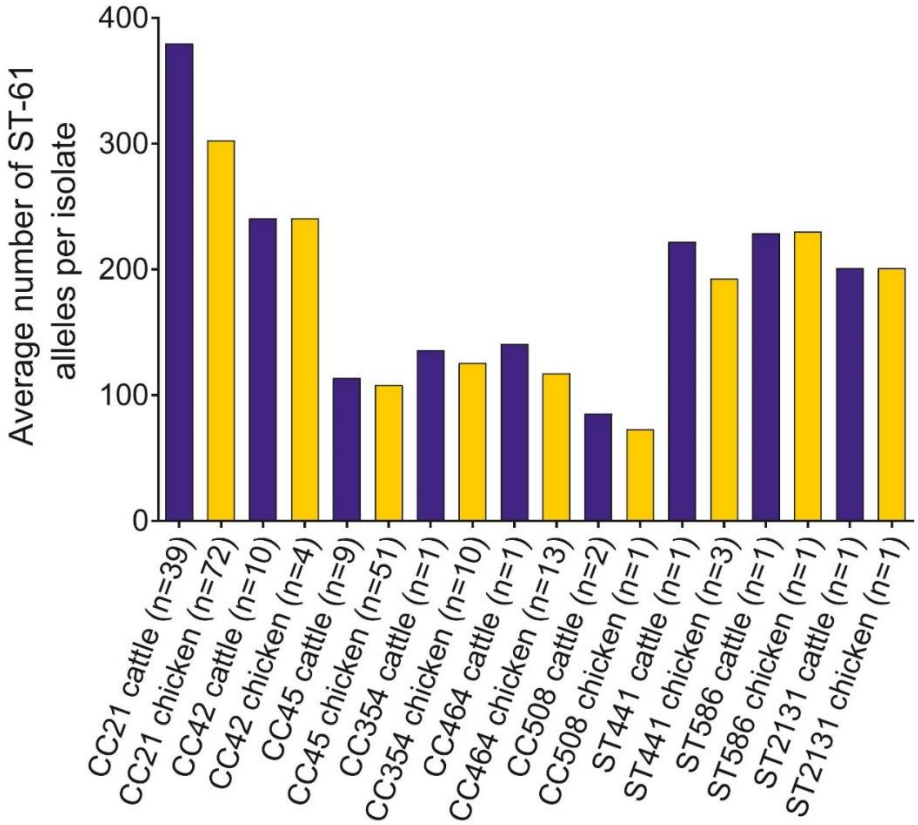


Figure 3.9. Prevalence of ST-61 complex alleles within the core genome of multiple lineages. The average number of ST-61 alleles per isolate (*y* axis) is calculated for all clonal complexes with isolates sampled from both cattle (blue) and chicken (yellow).

A more detailed analysis of homologous recombination across the genome was carried out using ClonalFrameML, to identify putative recombinant sites associated with the emergence of cattle specialist *C. jejuni.* Recombination had a considerable role in generating genetic diversity within the ST-61 clonal complex leading to 12.6 times as many nucleotide substitutions as mutation, compared to a recombination to mutation ratio (*r/m*) of 7.5 for the generalist ST-21 complex and 9.6 for the entire dataset (Table 3.2). These estimates are higher than the most widely cited *r/m* estimate for *C. jejuni* (2.2) (Vos and Didelot, 2009), that is based upon 7 MLST genes (Wilson et al., 2009).

Table 3.2. Recombination and transition/transversion parameters. Values are calculated by PhyML and ClonalFrameML.

Clonal complex	n	transition/ transversion	R/θ	1/δ	v	r/m = R/θ*δ*v
ST-21 complex	49	8.074	0.620008	0.00246347	0.0297719	7.493014396
ST-61 complex	50	4.295	0.590489	0.00216759	0.0462289	12.59355179
ST-21 / ST-61 subset used in ClonalFrameML	99	8.093	0.606459	0.00209515	0.028986	8.390244409
Whole dataset	1198	-	0.644303	0.00127492	0.0189957	9.599807437

However, for recombining bacteria, *r/m* estimates can vary widely even among lineages within species (Croucher et al., 2013; Didelot et al., 2012). Furthermore, recombination is likely to play a lesser role on the highly conserved MLST loci than it does throughout the genome. Results of this chapter suggest that recombination is associated with the emergence of the ST-61 complex, and the process of cattle specialization. A total of 5019 recombination events were recorded (Appendix, Figure S3.2), of which 175 occurred on the tree branch separating ST-61 from ST-21 complex isolates. These recombination events mapped to 227 genes including 42 whole gene replacements and 185 mosaic genes (Appendix, Table S3.2). Unconfirmed clues about the possible adaptive role of recombination come from the inferred function of recombinant genes. A total of 29.8% were related to metabolic functions, 10.5% were related to cell envelope biogenesis, and 22.9% were known flagella genes linked to cell motility (Appendix, Table S3.2). Alterations to flagella are known to be important for colonization of the host environment (Chen et al., 2011), particularly in *C. jejuni* and *Helicobacter pylori* where flagellar motility is a key adaptive trait enabling bacteria to move effectively through viscous environments such as the lumen and mucus layer of the gastrointestinal tract (Celli et al., 2009). Similar analysis of ST-42 complex inferred 184 recombination events, mapping to

621 genes, 102 of which were also found to have recombined on the branch leading to the ST-61 complex (Appendix, Table S3.3).

Convergent evolution reveals adaptive genomic changes

It can be challenging to differentiate adaptive genomic changes, that provide an advantage in the host, from changes resulting from genetic bottlenecks and drift (Sheppard et al., 2018). Among the most compelling evidence for genome adaptation occurs when divergent lineages independently acquire convergent genetic changes, which are not present in their common ancestor (homoplasy). Preliminary automated analysis of homoplasy (Crispell et al., 2019) identified variation in the number of homoplasies among clonal complexes (Figure 3.10). A more detailed analysis, consistent with studies identifying chicken adaptation in *Staphylococcus aureus* (Murray et al., 2017), was conducted, to identify homoplasy between phylogenetically distant cattle specialist ST-61 and ST-42 complexes.

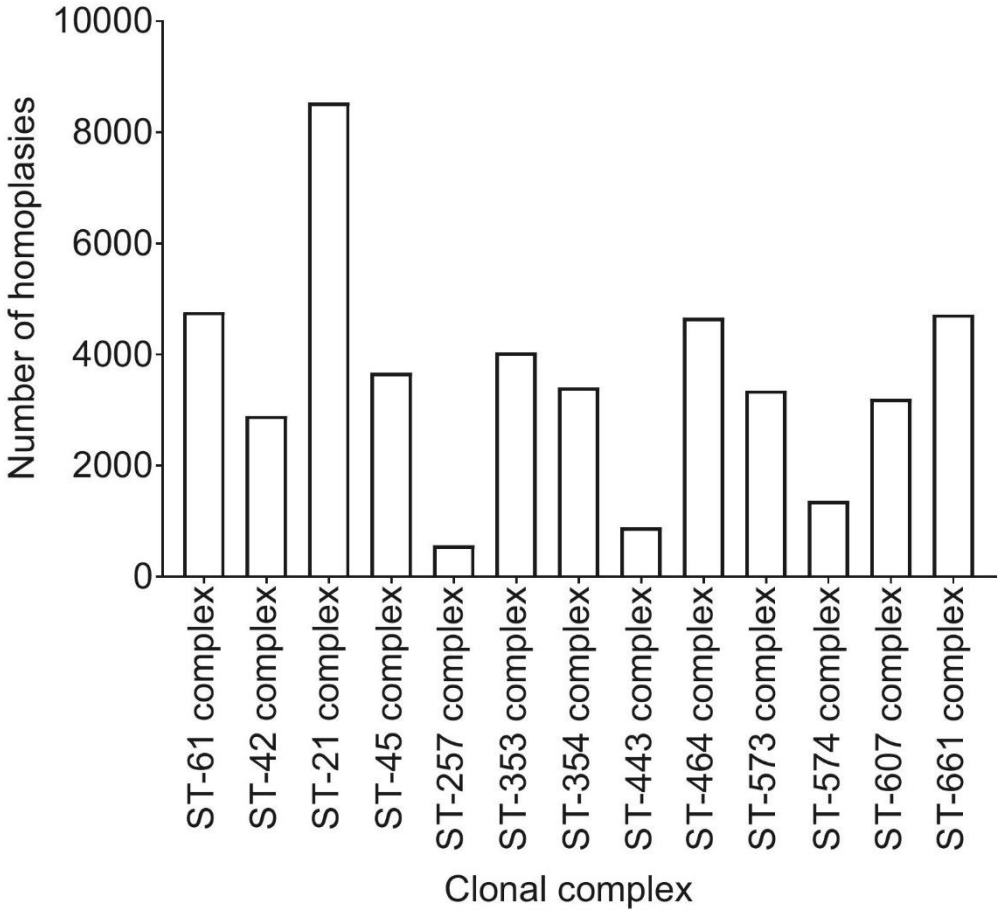


Figure 3.10. Homoplasies identified in 13 *C. jejuni* clonal complexes. Number of homoplasies identified using HomoplasyFinder is shown for 13 clonal complexes.

Hypothesizing that differential gene prevalence in host associated lineages is evidence of adaptation, 326 genes that were statistically associated with either ST-61 or ST-21 complexes (Fisher's exact test, *p <0.0005*), were first identified. Focussing on genes that were present at >90% in one complex and <10% in the other, 35 genes that were over-represented (Fisher's exact test, p<10e-13) in the ST-61 complex and 39 genes that were largely absent, were identified. Additionally, 175 inferred recombination blocks (mapping to 227 genes) were identified on the branch leading to the ST-61 complex. The presence/absence of these elements was investigated in ST-42 and 12 other clonal complexes to identify homoplasy (Appendix, Table S3.2).

Of the ST-61 complex-associated genes (n=35) and recombination blocks (n=175), 6 genes and 9 recombinant alleles (respectively) that were also present in ST-42 isolates (Figure 3.11), and were largely absent in ST-21 complex isolates and chicken-specialist ST-257, ST-353, ST-443, ST-573, ST-574, ST-607 and ST-661 complexes (Table 3.1, Figure 3.12, Figure 3.13), were identified.

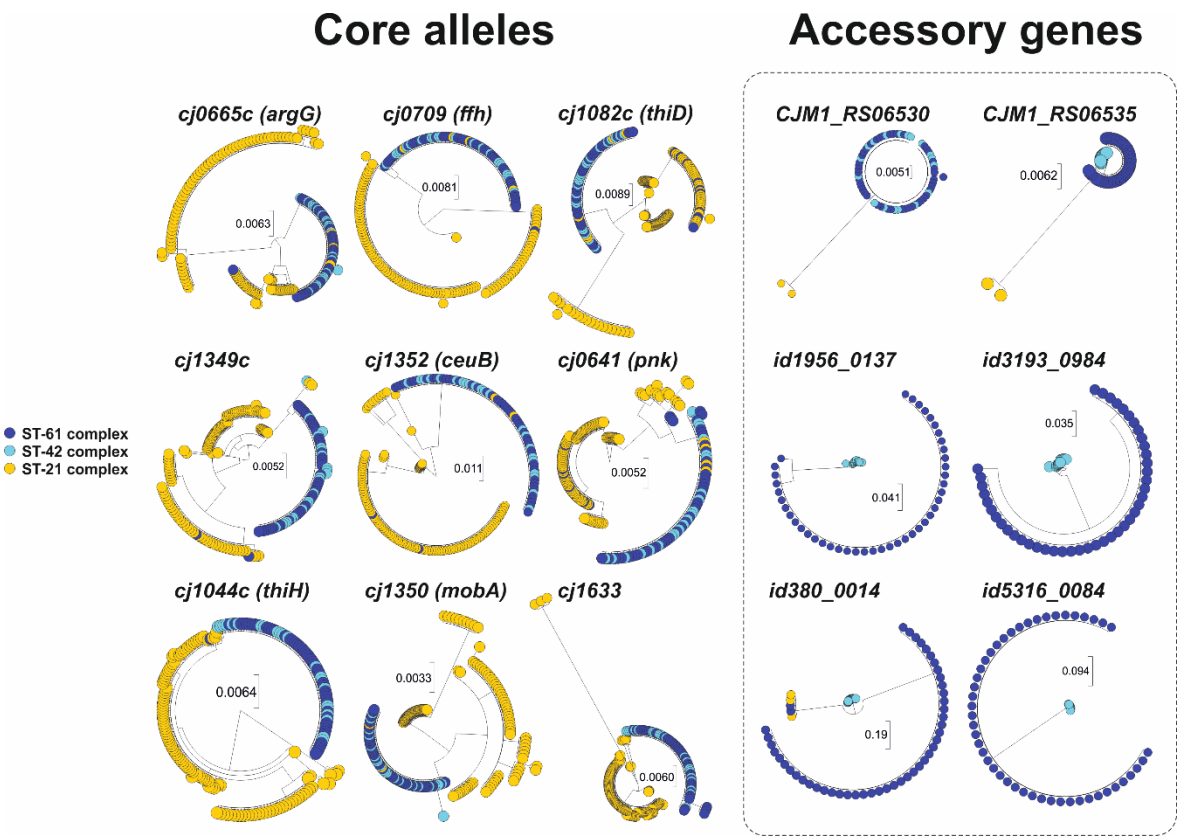


Figure 3.11. Homoplasy genes between ST-61 and ST-42 complexes. Single-gene trees for 9 candidate cattle-adaptive genes demonstrating homoplasy among cattle specialist lineages. Trees show intermingled clusters of isolates from the two cattle specialist lineages ST-61 (blue) and ST-42 (light blue), separated from the generalist lineage ST-21 yellow).

Accessory genome homoplasy included homologues of genes *cj1340* and *cj1341* from the *maf* gene family encoding flagella biosynthesis and bacterial motility factor proteins (Guerry et al., 2006; Novik et al., 2010). Recombinant genes with homoplasious alleles included 6 genes (*pnk, argG, thiH, thiD, mobA, ceuB*) associated with metabolism or transport, including genes involved in thiamine biosynthesis (*thiH* and *thiD*), molybdenum cofactor biosynthesis (*mobA*) for respiratory enzymes, and a putative fibronectin/fibrinogen-binding protein gene (*cj1349*) involved in host cell adhesion (Table 3.1, Figure 3.14).

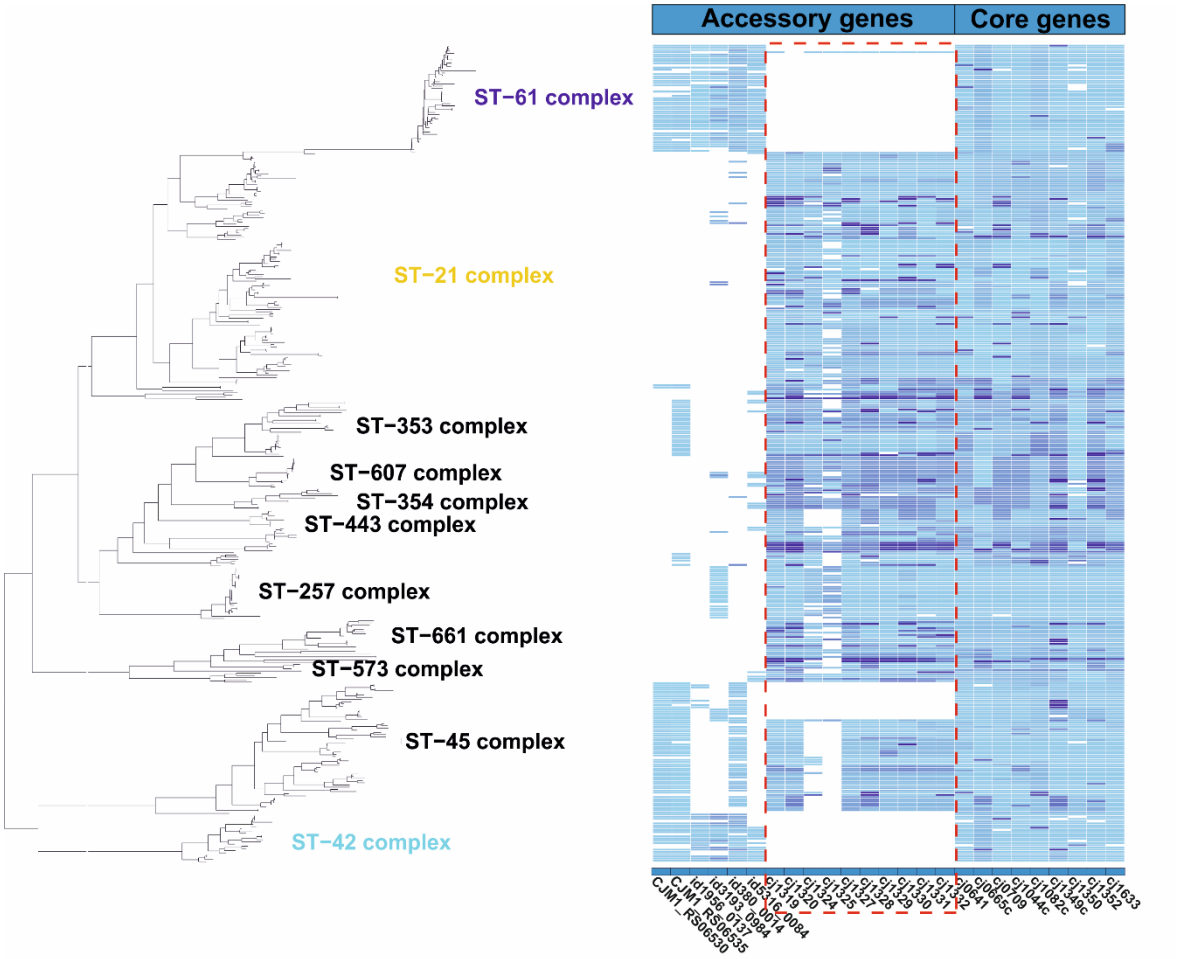


Figure 3.12. Phylogenetic structure of 381 *C. jejuni* isolates of cattle, sheep and chicken isolates and the distribution of the candidate cattle-associated genes and their alleles. Left panel: All genome sequences were from ST-61 (blue), ST-21 (yellow), ST-45 (black), ST-42 (light blue), ST-257 (black), ST-353 (black), ST-354 (black), ST-443 (black), ST-464, ST-573 (black), ST-574, ST-607 (black) and ST-661 (black) complexes. Right panel: Each bar indicates the presence of a gene for each isolate while the difference in colour corresponds to different allele of the same gene. The red box highlights a block of genes absent in ST-61 and ST-42 complexes. The phylogenetic trees were reconstructed using an approximation of the maximum-likelihood algorithm in RAxML.

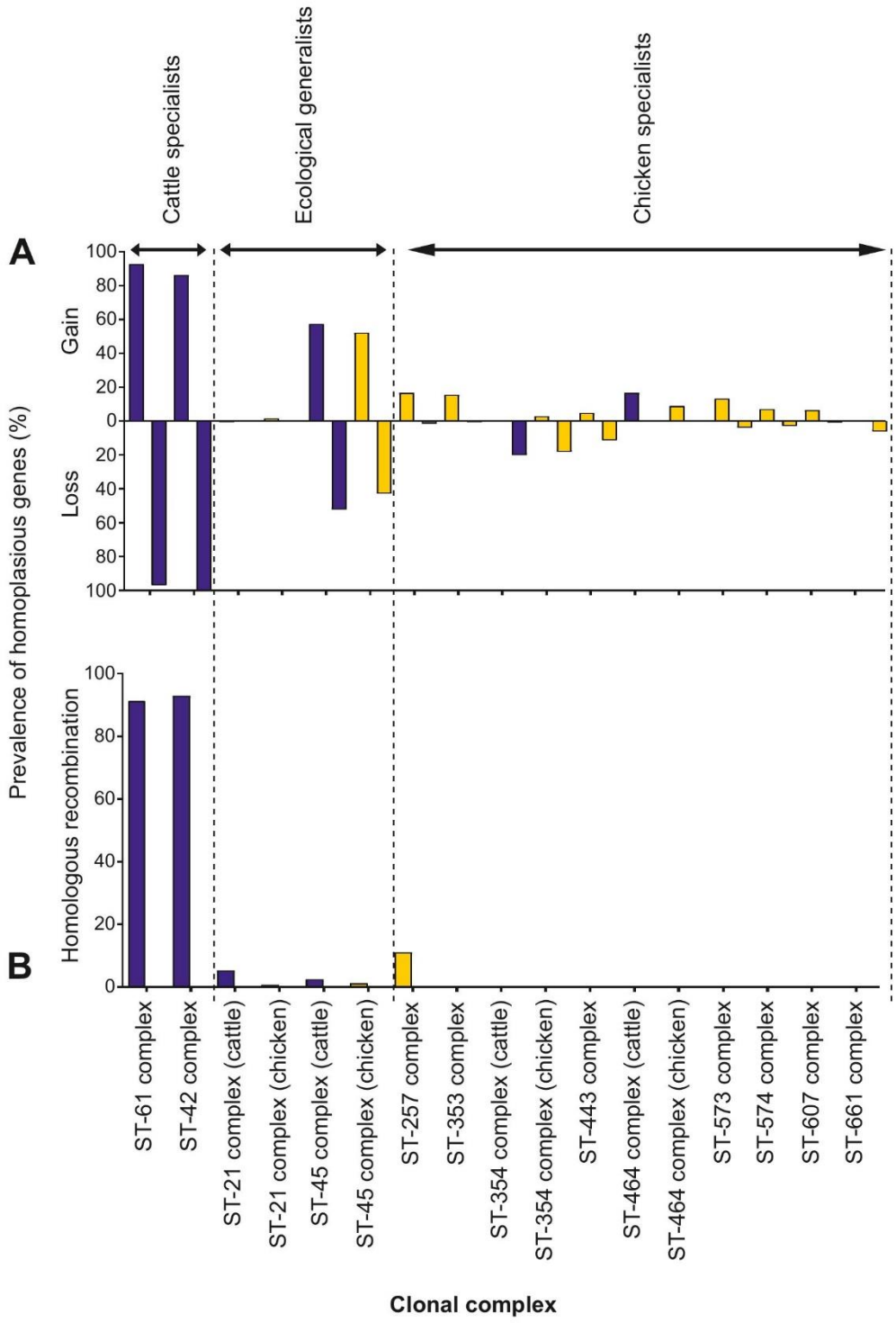


Figure 3.13. Prevalence of ST-61-42 complex homoplasy in 13 *C. jejuni* clonal complexes. (A) Prevalence of homoplasious genes gained (upper plot) and that are lost (bottom plot) in ST-61 and in 12 other *C. jejuni* ST-complexes with isolates sampled from both cattle (blue) and chicken (yellow). (B) Prevalence of homoplasious recombination in ST-61 and in 12 other *C. jejuni* ST-complexes with isolates sampled from both cattle (blue) and chicken (yellow). Most homoplasies were largely restricted to the cattle specialist ST-61 and ST-42 complexes.

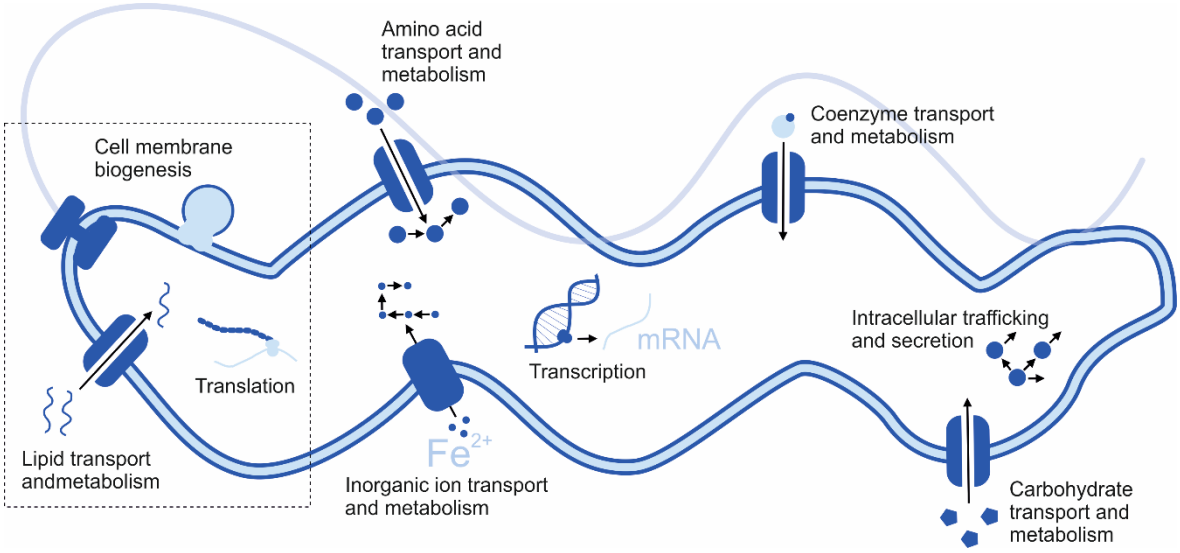


Figure 3.14. Summary of cattle adaptation based on comparative genomics indicating biological pathways associated with deletion of the glycosylation gene block.

Accessory gene loss is linked to cattle adaptation

A striking finding was the identification of gene blocks, including 10 genes, that were largely missing in ST-61 (n=48/49) and ST-42 (n=22/22) complex isolates, but present in all ST-21 and other chicken-associated complexes (Table 3.1, Figure 3.12, Figure 3.13, Appendix, Table S3.2). These genes were co-located in the genome, organised into 2 operons - suggesting functional linkage (Table 3.1), and included *cj1319, cj1320, cj1324*, *cj1325*, *cj1327 (neuB2)*, *cj1328 (neuC2)*, *cj1329*, *cj1330*, *cj1331 (ptmB)* and *cj1332 (ptmA),* all located in a region encoding the flagellin O-linked glycosylation system. Genes *cj1324* and *cj1325* encode enzymes necessary for the biosynthesis of legionaminic acid, a surface exposed polysaccharide that has been shown in previous studies to be essential for chicken colonization (Howard et al., 2009) (Table 3.1). Furthermore genes *cj1331 (ptmB)* and *cj1332 (ptmA)* have been shown to be involved in post translational modification of the flagellin protein first described in *C. coli* VC167 strain (Table 3.1) (Guerry et al., 1996). The independent loss of these genes since the emergence of the two major cattle specialist lineages is not consistent with a random bottlenecking event associated with colonization, and may indicate selection against strains that carry superfluous genes (beneficial in chicken and not cattle) because of a fitness cost to the bacterium (Morley et al., 2015; Koskiniemi et al., 2012; Kettler et al., 2007). Similar analysis for ST-42 complex revealed the loss of 34 genes all of which were also lost in ST-61 complex (Appendix, Table S3.3) but *in vivo* colonization assays would be necessary to confirm the adaptive advantage of gene loss.

Cattle adaptation is associated with functional variation in cell hydrophobicity, autoagglutination and biofilm formation

Identification of homplasious genomic signatures provides broad information about adaptive evolution but focusing on specific groups of genes, such as the lost glycosylation gene block, provides information about the functional basis of phenotypic adaptation. To investigate this, phenotype assays were conducted *in vitro* on 36 *C. jejuni* strains from cattle specialist (ST-61 and ST-42 complex), generalist (ST-21 complex) and chicken specialist lineages, where the glycosylation gene block was differentially present, and two engineered ΔGB (gene block) mutant strains (ST-661 and ST-257 complex) where the glycosylation gene block (*cj1324-1332*) was specifically deleted by allelic exchange mutagenesis. [N.B. The mutant strains were engineered by Aidan Taylor, University of Sheffield, Sheffield, UK.] [N.B. The phenotype assays were performed in collaboration with Aidan Taylor, University of Sheffield, Sheffield, UK.]

Post-translational modifications of flagellin are known to be important in autoagglutination and hydrophobicity (Guerry et al., 2006; Misawa and Blaser, 2000) and this is associated with virulence in gram-negative bacteria (Nothaft and Szymanski, 2010), including *Neisseria gonorrhoeae* (Swanson et al., 1971) and *Bordetella pertussis* (Menozzi et al., 1994). The engineered ΔGB mutant strains both displayed a dramatic loss of hydrophobicity, measured as an increase in the salt concentration at which cell aggregation occurs, compared to the otherwise isogenic wild-type parental strains (Figure 3.15A). Furthermore, *C. jejuni* strains from cattle specialist lineages displayed a significant reduction in hydrophobicity compared to those from chicken specialist lineages (Mann-Whitney *U*, P = 0.013) (Figure 3.15B). In the generalist ST-21 lineage, isolates from cattle showed a trend towards reduced hydrophobicity compared to those from chicken, though not significant (Figure 3.15B). Similarly, autoagglutination was significantly reduced in the engineered ΔGB mutant strains compared to the isogenic wild-type parents (unpaired t-test, P = 0.0343 and P = 0.0142) (Figure 3.15E).

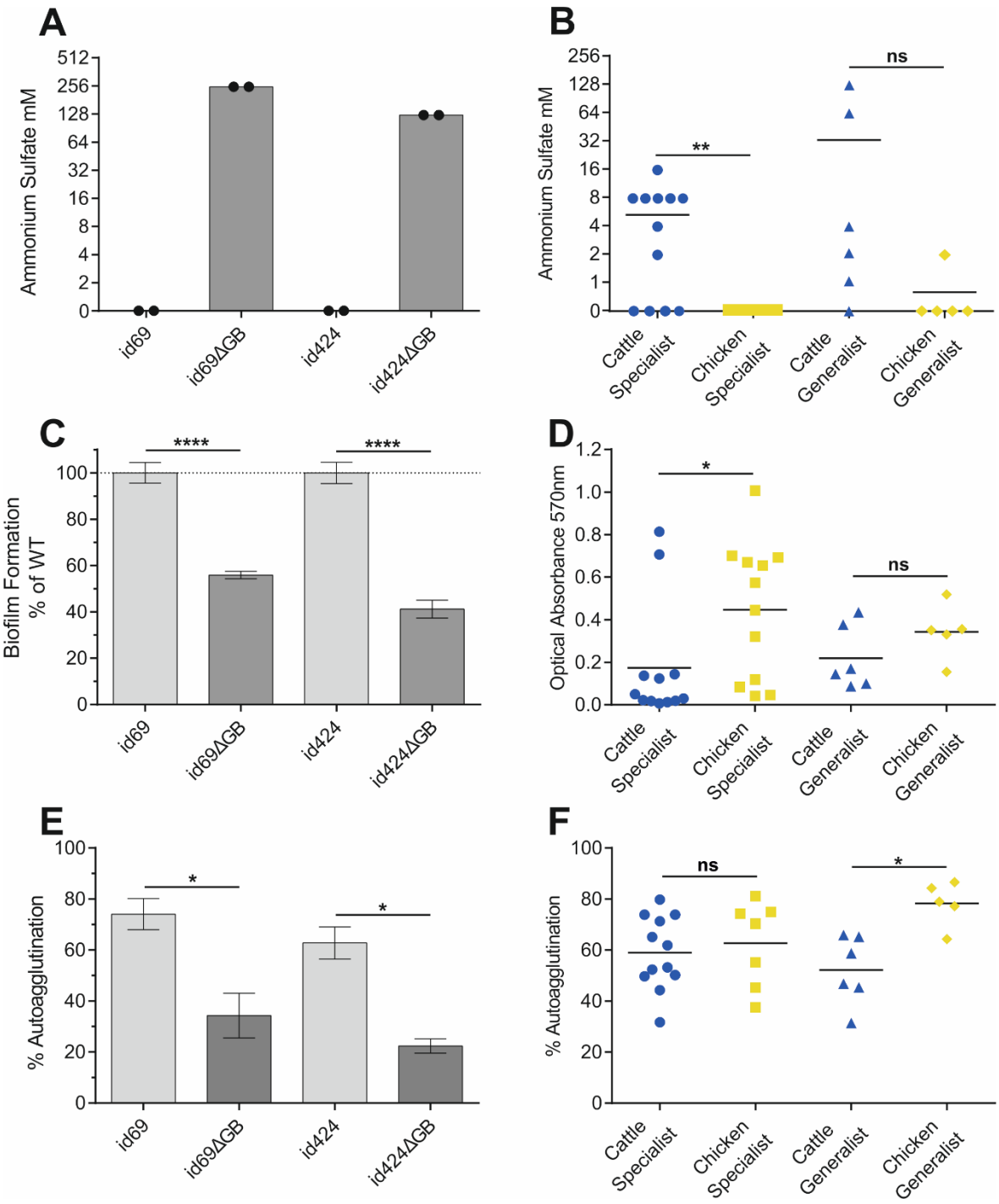


Figure 3.15. Evidence of cattle associated phenotype variation in *C. jejuni*. (A-F) Phenotype assays comparing two wild-type chicken specialist strains, id69 and id424 (Appendix, Table S3.1) and isogenic glycosylation gene block deletion mutants (id69ΔGB and id424ΔGB), and natural specialist and generalist strains from cattle and chicken. (A, C, E) Chicken associated lineage mutants show a marked decrease in cellular hydrophobicity, biofilm formation and autoagglutination compared to wild-type strains. In assays of natural strain collections, cattle specialist strains show (B) significantly greater mean ammonium sulfate concentration (mM) indicating decreased cell hydrophobicity, and (D) significantly decreased ability to from biofilms measured as optical absorbance (at 570 nm). (F) No difference was observed in autoagglutination in specialist strains from cattle and chicken but generalist strains from cattle showed decreased ability to autoagglutinate. Significance was tested using Mann-Whitney *U* test with * indicating $P < 0.05$, ** $P < 0.01$ and **** $P < 0.001$. The horizontal line in each plot represents the mean value.

A significant reduction was also observed among ST-21 complex generalist isolates from cattle compared to those from chickens (Mann-Whitney *U*, P = 0.0173) and a non-significant reduction was observed in cattle specialist strains, compared to the chicken specialists (Figure 3.15F).

Biofilm formation has also been strongly linked with flagellar properties and function (Joshua, 2006). Consistent with this, a highly significant reduction in biofilm formation was observed in the ΔGB mutants compared to the wild-type parental strains (unpaired t-test, P < 0.0001) (Figure 3.15C). Similarly, chicken specialist isolates displayed enhanced biofilm formation compared to cattle specialist strains (Mann-Whitney *U*, P = 0.0387) (Figure 3.15D). For the generalist ST-21 lineage, strains originating from chicken showed a higher average ability to form biofilm compared to strains from cattle, though the difference was not significant (Figure 3.15D). There was no significant difference in biofilm formation between strains of ST-61 and ST-42 cattle specialist lineages.

Discussion

The well-established link between anthropogenic environmental change and the emergence of zoonoses usually focusses on pathogen spill-over and spread in new host populations (McMichael, 2004; Woolhouse et al., 2005). However, for multi-host pathogens, changes to the host niche can also have a major impact on pre-existing transmission networks, enhancing the risk of human infection. Undoubtedly the most dramatic change to the natural host niche of *C. jejuni* has occurred with the advent of intensive livestock production favouring sub-lineages associated with agricultural animal hosts, such as cattle (Sheppard et al., 2013; Morley et al., 2015; Thépault et al., 2017). This affords the opportunity to investigate the genomics and time-scale of host adaptation in the industrialised world's most common food-borne bacterial pathogen.

Host generalist *C. jejuni* lineages, containing isolates sampled from both chicken and cattle, are indicative of multiple recent host transition events (Figure 3.1). The work on this chapter demonstrates that against this dynamic backdrop of mingling bacterial populations, a cattle-associated lineage - the ST-61 complex - arose and proliferated. This process of host specialisation, in apparent sympatry, may have proceeded with the gradual step-wise differentiation from the generalist ST-21

complex ancestor, as exemplified by the existence of rare intermediate genotypes (Figure 3.2), ultimately leading to the emergence of successful endemic and epidemic clones circulating globally within cattle. In comparative analysis of the cattle-associated ST-42 complex, emergence from ecological generalist ancestors could not be confirmed.

Dating the emergence of cattle specialist *C. jejuni* provides clues about the influence of livestock intensification on pathogen evolution. Cattle were first domesticated from wild aurochs (*Bos primigenius*) in the early Neolithic age (McTavish et al., 2013). Based upon its distribution in wild ungulates (Navarro-Gonzalez et al., 2014; Díaz-Sánchez et al., 2013) it is likely that *Campylobacter* was present in some cattle throughout the 10,000 year history of domestication. However, the emergence of the modern cattle specialist ST-61 lineage was dated much more recently, in the mid-to-late 1800's, with most lineages emerging in the 20th century. This is consistent with the rapid evolution of cattle specialism coinciding with the industrial agriculture revolution and the intensification of animal production (Green, 1990; Fraser, 2008). Between 1820 and 1975, the number of cattle increased 8-fold across the world (FAO, n.d.; Fraser, 2005; Armbruster, 2004). New farming techniques, improved livestock breeding, and higher stocking densities led to amplified food production, but the host population explosion and global transmission networks have also fundamentally changed the cattle niche, potentially favouring strains that are adapted to specialist lineages.

Horizontal gene transfer (HGT) plays an important role in the evolution of *Campylobacter,* generating genetic diversity at twice the rate of *de novo* mutation (Wilson et al., 2009) and potentially conferring novel function on the recipient genome. Characterization of segments of DNA that have recombined in cattle specialists provided information about the host-specific gene pool. Furthermore, to differentiate host-adaptive genetic signatures from host-associations, that may reflect bottlenecking and drift from the ancestral gene pool, the analysis focussed on homoplasious genomic changes that occurred independently in divergent cattle specialist lineages. The analysis revealed combinations of genes and alleles that potentially promote the host-adaptive evolution and specialization in divergent lineages.

Cattle and poultry display clear differences in anatomy, physiology and metabolism. Consistent with this, candidate cattle-adaptive genetic variation occurred in genes involved in diverse functions. For example, the putative fibronectin/fibrinogen-binding protein gene *cj1349c* is important for *in vitro* adherence to chicken epithelial cells (Flanagan et al., 2009), therefore anatomic and histological differences with cattle intestinal epithelial cells may explain the conservation of alleles in cattle specialist *C. jejuni*. Furthermore, differences in the diet of cattle and chicken may account for signatures of cattle adaptation in the genome. Vitamin B1 supplements are commonly given in poultry feed at high doses (Braunlich and Zintzen, 1979) while comparable host requirements are synthesized naturally in cattle by rumen bacteria, where the majority is absorbed in the small intestine with only low amounts reaching the large intestine (Santschi et al., 2005). This potentially means that *C. jejuni* in the cattle large intestine have distinct thiamine synthesis requirements accounting for homoplasy among thiamine biosynthesis genes (*thiH, thiD, thiE*) in cattle specialist *C. jejuni*. Similarly, the conservation of the *mobA* gene that encodes a key molybdopterin-guanine dinucleotide (MGD) biosynthesis protein that allows *C. jejuni* to use a range of alternative electron donors and acceptors for respiration, may help balance the rate of biosynthesis with the varying supply of potentially toxic (to cattle) molybdenum (Blakley, 2016; Suttle, 1991). While speculations about differences in host anatomy and physiology provide a context for considering the putative function of cattle association genetic variation, *in vivo* growth assays would be necessary to confirm adaptation.

The emergence of cattle specialist *C. jejuni* was also associated with significant gene loss. While reductive evolution is widely observed among bacteria within discrete niches (Batut et al., 2014), in most cases it is not directly adaptive but results from drift or linkage to other beneficial mutations (Sheppard et al., 2018). However, in *Campylobacter* and *Salmonella enterica,* loss of gene function has been shown to be a characteristic of some host-restricted lineages (Langridge et al., 2015; Koskiniemi et al., 2012; Morley et al., 2015; Kettler et al., 2007). Consistent with this, distinct gene loss signatures in the flagellin O-linked glycosylation system in cattle-associated *C. jejuni,* including specialist lineages, were observed. This was associated with reductions in cell hydrophobicity, autoagglutination and biofilm formation in comparisons among wild-type cattle and chicken specialist lineages and confirmed by studies with deletion mutants as

controls. The flagellin protein is post-translationally modified on the surface of the flagellum which is exposed to the extracellular/host environment and this is known to be associated with serospecificity (Logan et al., 2002), with O-linked glycosylation system genes over-represented in isolates from chicken compared to cattle (Howard et al., 2009). The absence of genes in this gene block, and the associated changes in cell surface charge and glycan moieties on the flagellum (Howard et al., 2009), may enable *C. jejuni* to attach better to the intestinal epithelium as changes in hydrophobicity in the flagella are believed to enable bacterial attachment to either hydrophobic or hydrophilic surfaces (Joshua, 2006). Furthermore, surface structures such as flagella are known to stimulate host innate and adaptive immunity (Jerome et al., 2011). Therefore, the loss of the glycosylation gene block, and associated phenotypic changes, may decrease recognition of strains by the host immune response. This provides a compelling basis for confirmatory *in vivo* experiments to investigate how gene loss may potentially influence proliferation in cattle.

Industrialized agriculture, including intensive livestock production, remains important to meet the nutritional requirements of a growing human population. However, the impact on livestock associated pathogens is seldom considered. Findings in this chapter suggest a change to the evolutionary history of *C. jejuni* that preceded the modern agricultural revolution. As the vast global cattle niche opened in the last century, a host generalist lifestyle ceased to be the only effective strategy for cattle colonization and thus specialist strains emerged - facilitated by HGT and significant gene gain and loss, related to differences in host diet, anatomy and histology - and were disseminated across the world. This highlights how the genomic plasticity of important zoonotic pathogens allows a response to radical anthropogenic changes to host ecology, potentially enhancing the risk of human infection. Further understanding the genetic and functional basis of host adaptation, particularly to livestock that constitute the majority of mammal biomass on earth, is important for the development of novel strategies, interventions and therapies to combat the increasing risk of pathogens with the capacity to spread from livestock to humans.

Materials and methods

Bacterial isolates

C. jejuni genome sequences were analysed for 1,198 isolates sampled from clinical cases of campylobacteriosis, chickens, ruminants, environmental sources, pets and wild birds from different countries (Appendix, Table S3.1) (Sheppard et al., 2013; Sheppard et al., 2014; Thépault et al., 2017; Morley et al., 2015). These included 1,065 from published studies, 119 isolate genomes available on the National Center for Biotechnology Information (NCBI) and 14 ST-61 complex isolates obtained from Animal and Plant Health Agency collections and sequenced as part of this chapter (Appendix, Table S3.1). These included isolates from cattle (n=8), sheep (n=3), chicken (n=1), giraffes (n=1) and humans (n=1), sampled between 1993 and 2003 (Appendix, Table S3.1). Comparative genomics centred on 101 ruminant isolates, principally cattle (n=93), and 1,097 isolates from other sources, representing multiple lineages defined by MLST. These included 26 and 10 isolates from the cattle associated ST-61 and ST-42 clonal complexes respectively, sampled principally from the UK (60%), Spain (12%) and the USA (10%) (Appendix, Table S3.1).

Culture, DNA extraction, and genome sequencing

C. jejuni strains were cultured on Columbia base agar plates containing 5% (v/v) horse blood and 5 µg ml^{-1} vancomycin in a MACS-VA500 workstation (Don Whitley Scientific Ltd, UK) under microaerobic conditions (10% v/v O_2, 5% v/v CO_2, 85% N_2, 42 °C). DNA was extracted using the QIAamp DNA Mini Kit (QIAGEN, Crawley, UK), according to manufacturer's instructions and quantified on a Nanodrop spectrophotometer prior to normalization and sequencing. High-throughput sequencing was performed using an Illumina MiSeq benchtop sequencer (Illumina, San Diego, CA). Short read paired-end data was assembled using SPAdes (version 3.10.035) and evaluated using QUAST (Gurevich et al., 2013). All assembled genomes were uploaded to a local instance of the BIGSdb web-based database platform (Jolley and Maiden, 2010) which allowed for archiving and gene-by-gene sequence alignment. A total of 14 genomes with assembled length >1.9Mbp, assembled in >500 contigs and with an N_{95}<800bp were considered of poor quality and excluded from the analyses.

Core and accessory genome variation and phylogenetic reconstruction

Sequence data was analysed using a reference pan-genome approach (Méric et al., 2014) in which a list was compiled for all of the genes present in: (i) reference

C. jejuni strains NCTC11168, 81116, 81-176, M1; (ii) plasmids pTet and pVir; (iii) annotations from the 1,198 genomes in this chapter. Closely related homologous genes were identified using BLAST (>70% sequence identity) and filtered out to produce a single gene list containing all the unique distinct genes for the dataset. Automatic annotations were obtained using RAST (Aziz et al., 2008) and a total of 3,855 unique genes were described for the dataset from 1,967,096 open reading frames (ORFs). Gene orthologues were aligned in a gene-by-gene manner (Méric et al., 2014) using MAFFT (Katoh, 2002) to produce a whole-genome multiple sequence alignment for all isolates of the dataset. The presence of individual gene sequences from the reference pan-genome list were detected in every genome of the dataset using BLAST with a match defined as >70% nucleotide identity over >50% of the gene length. This approach generated a gene presence/absence matrix summarising the presence and allelic variation of every gene in every genome. The core genome was defined as genes shared by all isolates, while a 'soft core' represented genes shared >95% by all isolates (Thépault et al., 2017). The remaining genes constituted the accessory genome. Phylogenetic trees in analysis of gene-by-gene alignments of core genes (Sheppard et al., 2013), single gene alignments, were reconstructed using the approximation of the maximum-likelihood algorithm implemented in RAxML v8.2.11 (Stamatakis, 2014) with the GTRGAMMA model. PhyML v3.3.2 (Guindon et al., 2010) and FastTree2 (Price et al., 2010) with GTR model of nucleotide substitution were also used to reconstruct the ML phylogenetic trees that were used for recombination and time-scaled analyses, respectively. Tajima's D was estimated over all sites which did not contain undetermined or missing bases using the PopGenome package (v2.6.1) in R (Pfeifer et al., 2014). Demographic reconstruction analysis was conducted using the non-parametric Bayesian Skyline model (Heled and Drummond, 2008) in BEAST2 (Bouckaert et al., 2019) and visualized with Tracer v1.7.1 (Rambaut et al., 2018). A GTR+G4 DNA substitution model was used in combination with a relaxed log-normal clock model and Coalescent Bayesian Skyline tree prior. A prior on the clock rate was set as a log-normal distribution with a mean value of 1×10^{-6} mutations per site per year (Duchêne et al., 2016).

Detection of recombination in *C. jejuni* ST-61 and ST-21 complexes

A subset of 99 isolates comprising 50 ST-61 complex isolates and 49 ST-21 complex isolates was used for further recombination analysis (Appendix, Table

S3.1). A reference pan-genome for this subset was created as described above. From a total of 174,730 ORFs, 2,570 genes remained after the removal of allelic variants. Of these, 1,498 genes were shared by 95% of all isolates, from which a gene-by-gene alignment was created using MAFFT (Katoh, 2002). Homologous recombination events were inferred using ClonalFrameML (Didelot and Wilson, 2015) using a guiding phylogeny reconstructed with PhyML (Guindon et al., 2010). The transition/transversion rate for ClonalFrameML was set to 8.093, a value calculated by PhyML. The resulting estimates were R/θ=0.606459 (relative rate of recombination to mutation), $1/\delta$=0.00209515 (inverse mean DNA import length) and v=0.028986 (mean divergence of imported DNA). The prevalence of exactly similar recombination regions was individually assessed in all isolates using BLAST with a threshold of 100% sequence identity. We adopted the conservative approach of excluding recombination regions with a sequence length >1000 bp, as these regions were likely to contain more than two genes and have reduced biological significance when using a concatenated gene-by-gene alignment as an input for this analysis. The relative number of substitutions introduced by recombination (r) and mutation (m) was calculated as the ratio $r/m = (R/\theta) \times \delta \times v$, with parameters inferred directly from ClonalFrameML using the transition/transversion ratio values and a guiding phylogeny computed with PhyML.

Detection of homoplasy in cattle adapted lineages

Homoplasy was defined as genetic similarity in divergent lineages that was absent in their common ancestor. Analysis was carried out to detect homoplasious accessory gene loss/gain and homologous recombination based on comparison of ST-61 complex isolates with ST-42, ST-21 and 11 other clonal complexes. Accessory gene presence was defined as a >70% BLAST match over 50% or more of any gene in the pan-genome list. For homologous recombination inferred in 49 ST-61 complex isolates using ClonalFrameML, recombinant sequence blocks (8 and 6191 bp) were extracted from the alignment and sequence <1000bp (n=140) was locally aligned using BLAST to 328 genomes. This reference genome set included isolates from cattle and chicken belonging to ST-21, ST-45, ST-42, ST-257, ST-353, ST-354, ST-443, ST-464, ST-573, ST-574, ST-607 and ST-661 complexes. An alignment match was considered homoplasious when there was an exact sequence match of 100% sequence identity and length found in two phylogenetically distinct lineages. Recombinant sequence >1000 bp in length

(n=35) were analysed using STRUCTURE software (Sheppard et al., 2009), to differentiate concatenates of different genes or gene fragments of genes in gene-by-gene alignments. Briefly, the No Admixture model with an independent allele frequency model was used to calculate the frequency of all alleles of 1,469 pan-genome loci that were present in the reference genome set. Analyses were performed with 100 burn-in cycles followed by 1000 iterations. Homoplasies were identified for 77 isolates (Appendix, Table S3.1) belonging to 13 clonal complexes using HomoplasyFinder (Crispell et al., 2019). Briefly, gene orthologues were aligned in a gene-by-gene manner using MAFFT (Katoh, 2002) to produce a whole-genome multiple sequence alignment. The core alignment (1,423 genes >90% present) was used to quantify homoplasies on branches leading to 13 ST-complexes using HomoplasyFinder (Crispell et al., 2019).

Time-scaled phylogenetic analyses

The timescale of emergence and diversification of the cattle associated ST-61 complex was estimated using 41 isolates with a known isolation date, with ST-5161 isolate as the outgroup. Core genes were aligned using MAFFT (Katoh, 2002). Genes were then ordered using *C. jejuni* NCTC11168 and concatenated to generate an ordered core gene alignment. This core gene alignment was used to reconstruct a phylogeny using RAxML (Stamatakis, 2014). Recombination regions, inferred by ClonalFrameML, were masked from the core genome sequence alignment (i.e. replaced with gaps) using a custom script (available on https://github.com/kwongj/cfml-maskrc) and the temporal signal was investigated with linear regression analysis of the root-to-tip distances against the sampling years using TempEst v1.5.1 (Rambaut et al., 2016). A time-scale phylogeny was constructed using BEAST2 v2.5.0 (Bouckaert et al., 2019), based on 1,320 variable sites, using the GTR +G4 model of DNA substitution. The marginal likelihood of 9 different combinations of the strict clock, relaxed clock exponential and relaxed clock log-normal models with the constant population, exponential population growth and Bayesian skyline models was compared, using generalized stepping-stone sampling (GSS) (Baele et al., 2016), run for 1 million iterations in 30 path steps. The relaxed log-normal clock with Bayesian skyline model showed the largest marginal likelihood. Input xml files were prepared using BEAUti2 v.2.5.0 (Bouckaert et al., 2019). A prior on the clock rate was set as a log-normal distribution with a mean value of 1×10^{-6} mutations per site per year (Duchêne et al., 2016). Markov chains

were run for 100 million generations, sampled every 10,000 generations with the first 10,000,000 generations (10%) discarded as burn-in. Three independent runs of 100 million generations were conducted and convergence was assessed by checking that the effective sample size (ESS) of all parameters exceeded 400 using Tracer v1.7.1 (Rambaut et al., 2018). TreeAnnotator (implemented in BEAST2 package) was used to generate a maximum clade credibility tree after discarding 10% burn-in.

Mutagenesis

The gene block absent from cattle specialists (*cj1324-1332*) was inactivated in two randomly selected chicken specialist strains from distinct sequence clusters id69 (ST-661) and id424 (ST-257) (Appendix, Table S3.4) by deletion of the majority of the open reading frames and replacement with a kanamycin resistance cassette through allelic exchange mutagenesis. The mutation vector was created using the Gibson isothermal assembly method as described previously (Taylor et al., 2017) using primers geneblockF1/R1 and F2/R2 to amplify upstream and downstream regions flanking the deletion and KanF and KanR primers to amplify the resistance cassette (Appendix, Table S3.4). The fragments were assembled into HincII digested pGEM3Zf to form pBLOCK. Competent cells of *C. jejuni* strains id69 and id424 were prepared by washing in cold 9% (w/v) sucrose, 15% (v/v) glycerol solution and were transformed with pBLOCK by electroporation. Mutant clones were selected by kanamycin resistance and correct insertion of the kanamycin cassette into the genome confirmed by PCR screening with primers geneblockF1 and geneblockR2 (Appendix, Table S3.4). The ΔGB (gene block) strains were screened in phenotype assays along with 36 *C. jejuni* isolates from cattle specialist ST-61 (n=6) and ST-42 (n=6) complexes, chicken specialist ST-257 (n=5), ST-353 (n=1), ST-573 (n=2), ST-574 (n=2), ST-607 (n=1) and ST-661 (n=1) complexes, as well as the generalist ST-21 (n=12) complex isolates of cattle, chicken and clinical origin (Appendix, Table S3.1). [N.B. The mutant strains were engineered by Aidan Taylor, University of Sheffield, Sheffield, UK.]

Cell surface hydrophobicity assay

Overnight growth was harvested from agar plates and resuspended in 2 mM Na Phosphate buffer (pH 7.4) to an optical density at 600 nm of approximately 1.0. Twenty microliters of cell suspension were aliquoted per well in U-bottomed 96 well

plates (Greiner BioOne 650161) and 180 µl of a two-fold serial dilution of ammonium sulfate (2 M to 0.98 mM) added. Plates were incubated statically at room temperature for 3 days. The minimum concentration of ammonium sulfate permitting aggregation of cells defines the point of hydrophobicity (Howard et al., 2009). [N.B. The cell hydrophobicity assay was performed in collaboration with Aidan Taylor, University of Sheffield, Sheffield, UK.]

Auto-agglutination assay

Overnight growth was harvested from agar plates and resuspended in phosphate buffered saline (PBS) to an optical density at 600 nm of approximately 1.0. An amount of 1 ml of the suspension was sampled and an accurate starting optical density was measured. Cell suspensions were then incubated under standard microaerobic conditions statically in plastic tubes for 1.5 hours. The top 1 ml of the suspension was then carefully removed, and the optical density measured. The percentage (%) of auto-agglutination was determined by subtracting the OD_{600} measured after 1.5 h from the starting OD_{600}, dividing by the starting OD_{600} and multiplying by 100 [(Starting OD_{600} – Final OD_{600})/Starting OD_{600}) x 100] (Howard et al., 2009). [N.B. The auto-agglutination assay was performed in collaboration with Aidan Taylor, University of Sheffield, Sheffield, UK.]

Biofilm assay

Overnight growth was harvested from agar plates and resuspended in Muller-Hinton broth to an optical density at 600 nm of approximately 0.2. Two hundred microliters of the cell suspension were aliquoted into 12 replicate wells of a tissue culture treated 96 well plate (CellStar 655180). Plates were incubated under standard microaerobic conditions statically for 3 days. The culture was removed, and wells stained with 1% (w/v) crystal violet in 90% (v/v) ethanol for 5 min. The stain was removed, and wells washed 3 times with dH_2O. Biofilms were apparent as clearly stained rings corresponding to the liquid-air interface. Plates were then allowed to air dry before de-staining with 300 µl of 80% v/v ethanol, 20% v/v acetone for 10 min. One hundred and fifty microliters were transferred per well to an optically clear 96 well plate and the optical density measured at 570 nm in a SpectraMax plate reader (Molecular Devices). Muller-Hinton broth controls were performed to normalise plate to plate variation. [N.B. The biofilm assay was performed in collaboration with Aidan Taylor, University of Sheffield, Sheffield, UK.]

References

Allen, R.C., 2005. English and Welsh Agriculture, 1300-1850: Output, Inputs, and Income. *Mimeo. Nuffield College, Oxford* [Online], pp.1–54.

Armbruster, K., 2004. Dominion: The Power of Man, the Suffering of Animals, and the Call to Mercy. *Interdisciplinary Studies in Literature and Environment*, 11(1), pp.235–236.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A. and Zagnitko, O., 2008. The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics*, 9(1), p.75.

Baele, G., Lemey, P. and Suchard, M.A., 2016. Genealogical Working Distributions for Bayesian Model Testing with Phylogenetic Uncertainty. *Systematic Biology*, 65(2), pp.250–264.

Bar-On, Y.M., Phillips, R. and Milo, R., 2018. The biomass distribution on Earth. *Proceedings of the National Academy of Sciences*, 115(25), pp.6506–6511.

Batut, B., Knibbe, C., Marais, G. and Daubin, V., 2014. Reductive genome evolution at both ends of the bacterial population size spectrum. *Nature Reviews Microbiology*, 12(12), pp.841–850.

Blakley, B.R., 2016. Molybdenum Poisoning. MSD Manual Veterinary Manual, Merck & Co. Inc. Kenilworth NJ USA.

Bollongino, R., Burger, J., Powell, A., Mashkour, M., Vigne, J.-D. and Thomas, M.G., 2012. Modern Taurine Cattle Descended from Small Number of Near-Eastern Founders. *Molecular Biology and Evolution*, 29(9), pp.2101–2104.

Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., Matschiner, M., Mendes, F.K., Müller, N.F., Ogilvie, H.A., du Plessis, L., Popinga, A., Rambaut, A., Rasmussen, D., Siveroni, I., Suchard, M.A., Wu, C.-H., Xie, D., Zhang, C., Stadler, T. and Drummond, A.J., 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLOS Computational Biology*, 15(4), p.e1006650.

Braunlich, K. and Zintzen, H., 1979. *Vitamin B1 in Animal Nutrition* [Online]. o.

Celli, J.P., Turner, B.S., Afdhal, N.H., Keates, S., Ghiran, I., Kelly, C.P., Ewoldt,

R.H., McKinley, G.H., So, P., Erramilli, S. and Bansil, R., 2009. Helicobacter pylori moves through mucus by reducing mucin viscoelasticity. *Proceedings of the National Academy of Sciences*, 106(34), pp.14321–14326.

Chen, S., Beeby, M., Murphy, G.E., Leadbetter, J.R., Hendrixson, D.R., Briegel, A., Li, Z., Shi, J., Tocheva, E.I., Müller, A., Dobro, M.J. and Jensen, G.J., 2011. Structural diversity of bacterial flagellar motors. *The EMBO Journal*, 30(14), pp.2972–2981.

Crispell, J., Balaz, D. and Gordon, S.V., 2019. HomoplasyFinder: a simple tool to identify homoplasies on a phylogeny. *Microbial Genomics*, 5(1).

Croucher, N.J., Finkelstein, J.A., Pelton, S.I., Mitchell, P.K., Lee, G.M., Parkhill, J., Bentley, S.D., Hanage, W.P. and Lipsitch, M., 2013. Population genomics of post-vaccine changes in pneumococcal epidemiology. *Nature Genetics*, 45(6), pp.656–663.

Cutler, S.J., Fooks, A.R. and van der Poel, W.H.M., 2010. Public Health Threat of New, Reemerging, and Neglected Zoonoses in the Industrialized World. *Emerging Infectious Diseases*, 16(1), pp.1–7.

Daszak, P., 2000. Emerging Infectious Diseases of Wildlife-- Threats to Biodiversity and Human Health. *Science*, 287(5452), pp.443–449.

Díaz-Sánchez, S., Sánchez, S., Herrera-León, S., Porrero, C., Blanco, J., Dahbi, G., Blanco, J.E., Mora, A., Mateo, R., Hanning, I. and Vidal, D., 2013. Prevalence of Shiga toxin-producing Escherichia coli, Salmonella spp. and Campylobacter spp. in large game animals intended for consumption: Relationship with management practices and livestock influence. *Veterinary Microbiology*, 163(3–4), pp.274–281.

Didelot, X., Eyre, D.W., Cule, M., Ip, C.L.C., Ansari, M., Griffiths, D., Vaughan, A., O'Connor, L., Golubchik, T., Batty, E.M., Piazza, P., Wilson, D.J., Bowden, R., Donnelly, P.J., Dingle, K.E., Wilcox, M., Walker, A., Crook, D.W., A Peto, T.E. and Harding, R.M., 2012. Microevolutionary analysis of Clostridium difficile genomes to investigate transmission. *Genome Biology*, 13(12), p.R118.

Didelot, X. and Wilson, D.J., 2015. ClonalFrameML: Efficient Inference of Recombination in Whole Bacterial Genomes. *PLoS Computational Biology*, 11(2), pp.1–18.

Dingle, K.E., Colles, F.M., Wareing, D.R.A., Ure, R., Fox, A.J., Bolton, F.E., Bootsma, H.J., Willems, R.J.L., Urwin, R. and Maiden, M.C.J., 2001. Multilocus Sequence Typing System for Campylobacter jejuni. *Journal of*

Clinical Microbiology, 39(1), pp.14–23.

Duchêne, S., Holt, K.E., Weill, F.-X., Le Hello, S., Hawkey, J., Edwards, D.J., Fourment, M. and Holmes, E.C., 2016. Genome-scale rates of evolutionary change in bacteria. *Microbial Genomics*, 2(11), p.e000094.

Eurosurveillance editorial team, 2015. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2013. *EFSA Journal*, 13(1), p.3991.

FAO, n.d. Production. Live Animals. Latest update: 07/05/2018. Accessed: 09/05/2018.

FAO, 2004. *FAOSTAT database* [Online]. Food and Agricultural Organization of the United Nations, Rome.

Flanagan, R.C., Neal-McKinney, J.M., Dhillon, A.S., Miller, W.G. and Konkel, M.E., 2009. Examination of Campylobacter jejuni Putative Adhesins Leads to the Identification of a New Protein, Designated FlpA, Required for Chicken Colonization. *Infection and Immunity*, 77(6), pp.2399–2407.

Fraser, D., 2005. Animal welfare and the intensification of animal production An alternative interpretation. *FAO* [Online].

Fraser, D., 2008. Animal Welfare and the Intensification of Animal Production. pp.167–189.

Friedman, C.R., Neiman, J., Wegener, H.C., T.R.V., 2000. Epidemiology of Campylobacter jejuni infections in the United States and other industrialised nations. *I. Nachamkin, M.J. Blaser (Eds), Campylobacter, ASM Press, Washington, D.C.* [Online], pp.121–138.

Frisch, M.B., Castillo-Ramírez, S., Petit, R.A., Farley, M.M., Ray, S.M., Albrecht, V.S., Limbago, B.M., Hernandez, J., See, I., Satola, S.W. and Read, T.D., 2018. Invasive Methicillin-Resistant Staphylococcus aureus USA500 Strains from the U.S. Emerging Infections Program Constitute Three Geographically Distinct Lineages. *mSphere*, 3(3), pp.1–14.

Graña-Miraglia, L., Lozano, L.F., Velázquez, C., Volkow-Fernández, P., Pérez-Oseguera, Á., Cevallos, M.A. and Castillo-Ramírez, S., 2017. Rapid Gene Turnover as a Significant Source of Genetic Variation in a Recently Seeded Population of a Healthcare-Associated Pathogen. *Frontiers in Microbiology*, 8(SEP), pp.1–11.

Green, B.H., 1990. Agricultural intensification and the loss of habitat, species and amenity in British grasslands: a review of historical change and assessment of

future prospects+. *Grass and Forage Science*, 45(4), pp.365–372.

Guerry, P., Doig, P., Alm, R. a, Burr, D.H., Kinsella, N. and Trust, T.J., 1996. Identification and characterization of genes required for post-translational modification of Campylobacter coli VC167 flagellin. *Molecular Microbiology*, 19(2), pp.369–378.

Guerry, P., Ewing, C.P., Schirm, M., Lorenzo, M., Kelly, J., Pattarini, D., Majam, G., Thibault, P. and Logan, S., 2006. Changes in flagellin glycosylation affect Campylobacter autoagglutination and virulence. *Molecular Microbiology*, 60(2), pp.299–311.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W. and Gascuel, O., 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology*, 59(3), pp.307–321.

Gurevich, A., Saveliev, V., Vyahhi, N. and Tesler, G., 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, 29(8), pp.1072–1075.

Heled, J. and Drummond, A.J., 2008. Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology*, 8(1), p.289.

Howard, S.L., Jagannathan, A., Soo, E.C., Hui, J.P.M., Aubry, A.J., Ahmed, I., Karlyshev, A., Kelly, J.F., Jones, M.A., Stevens, M.P., Logan, S.M. and Wren, B.W., 2009. Campylobacter jejuni Glycosylation Island Important in Cell Charge, Legionaminic Acid Biosynthesis, and Colonization of Chickens. *Infection and Immunity*, 77(6), pp.2544–2556.

Jerome, J.P., Bell, J.A., Plovanich-Jones, A.E., Barrick, J.E., Brown, C.T. and Mansfield, L.S., 2011. Standing Genetic Variation in Contingency Loci Drives the Rapid Adaptation of Campylobacter jejuni to a Novel Host. *PLoS ONE*, 6(1), p.e16399.

Jolley, K.A. and Maiden, M.C., 2010. BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics*, 11(1), p.595.

Jones, B.A., Grace, D., Kock, R., Alonso, S., Rushton, J., Said, M.Y., McKeever, D., Mutua, F., Young, J., McDermott, J. and Pfeiffer, D.U., 2013. Zoonosis emergence linked to agricultural intensification and environmental change. *Proceedings of the National Academy of Sciences*, 110(21), pp.8399–8404.

Joshua, G.W.P., 2006. Biofilm formation in Campylobacter jejuni. *Microbiology*, 152(2), pp.387–396.

Kaakoush, N.O., Castaño-Rodríguez, N., Mitchell, H.M. and Man, S.M., 2015. Global Epidemiology of Campylobacter Infection. *Clinical Microbiology Reviews*, 28(3), pp.687–720.

Katoh, K., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30(14), pp.3059–3066.

Kettler, G.C., Martiny, A.C., Huang, K., Zucker, J., Coleman, M.L., Rodrigue, S., Chen, F., Lapidus, A., Ferriera, S., Johnson, J., Steglich, C., Church, G.M., Richardson, P. and Chisholm, S.W., 2007. Patterns and Implications of Gene Gain and Loss in the Evolution of Prochlorococcus. *PLoS Genetics*, 3(12), p.e231.

Koskiniemi, S., Sun, S., Berg, O.G. and Andersson, D.I., 2012. Selection-Driven Gene Loss in Bacteria. *PLoS Genetics*, 8(6), p.e1002787.

Kwan, P.S.L., Barrigas, M., Bolton, F.J., French, N.P., Gowland, P., Kemp, R., Leatherbarrow, H., Upton, M. and Fox, A.J., 2008. Molecular Epidemiology of Campylobacter jejuni Populations in Dairy Cattle, Wildlife, and the Environment in a Farmland Area. *Applied and Environmental Microbiology*, 74(16), pp.5130–5138.

Langridge, G.C., Fookes, M., Connor, T.R., Feltwell, T., Feasey, N., Parsons, B.N., Seth-Smith, H.M.B., Barquist, L., Stedman, A., Humphrey, T., Wigley, P., Peters, S.E., Maskell, D.J., Corander, J., Chabalgoity, J.A., Barrow, P., Parkhill, J., Dougan, G. and Thomson, N.R., 2015. Patterns of genome evolution that have accompanied host adaptation in Salmonella. *Proceedings of the National Academy of Sciences*, 112(3), pp.863–868.

Llarena, A.-K., Zhang, J., Vehkala, M., Välimäki, N., Hakkinen, M., Hänninen, M.-L., Roasto, M., Mäesaar, M., Taboada, E., Barker, D., Garofolo, G., Cammà, C., Di Giannatale, E., Corander, J. and Rossi, M., 2016. Monomorphic genotypes within a generalist lineage of Campylobacter jejuni show signs of global dispersion. *Microbial Genomics*, 2(10), p.e000088.

Logan, S.M., Kelly, J.F., Thibault, P., Ewing, C.P. and Guerry, P., 2002. Structural heterogeneity of carbohydrate modifications affects serospecificity of Campylobacter flagellins. *Molecular Microbiology*, 46(2), pp.587–597.

Manyi-Loh, C., Mamphweli, S., Meyer, E., Makaka, G., Simon, M. and Okoh, A., 2016. An Overview of the Control of Bacterial Pathogens in Cattle Manure. *International Journal of Environmental Research and Public Health*, 13(9),

p.843.

Maue, A.C., Mohawk, K.L., Giles, D.K., Poly, F., Ewing, C.P., Jiao, Y., Lee, G., Ma, Z., Monteiro, M.A., Hill, C.L., Ferderber, J.S., Porter, C.K., Trent, M.S. and Guerry, P., 2013. The Polysaccharide Capsule of Campylobacter jejuni Modulates the Host Immune Response. *Infection and Immunity*, 81(3), pp.665–672.

McMichael, A.J., 2004. Environmental and social influences on emerging infectious diseases: past, present and future. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359(1447), pp.1049–1058.

McTavish, E.J., Decker, J.E., Schnabel, R.D., Taylor, J.F. and Hillis, D.M., 2013. New World cattle show ancestry from multiple independent domestication events. *Proceedings of the National Academy of Sciences*, 110(15), pp.E1398–E1406.

Menozzi, F.D., Boucher, P.E., Riveau, G., Gantiez, C. and Locht, C., 1994. Surface-associated filamentous hemagglutinin induces autoagglutination of Bordetella pertussis. *Infection and immunity*, 62(10), pp.4261–9.

Méric, G., Yahara, K., Mageiros, L., Pascoe, B., Maiden, M.C.J., Jolley, K.A. and Sheppard, S.K., 2014. A reference pan-genome approach to comparative bacterial genomics: identification of novel epidemiological markers in pathogenic Campylobacter. *PloS one*, 9(3), p.e92798.

Mira, A., Ochman, H. and Moran, N.A., 2001. Deletional bias and the evolution of bacterial genomes. *Trends in Genetics*, 17(10), pp.589–596.

Misawa, N. and Blaser, M.J., 2000. Detection and Characterization of Autoagglutination Activity by Campylobacter jejuni. *Infection and Immunity*, 68(11), pp.6168–6175.

Mitchell, B.R., 1998. *International historical statistics : the Americas 1750-1988, Europe 1750-1993, Africa, Asia & Oceania 1750-1993* [Online]. 2nd ed. New York, Stockton Press, London, Macmillan Reference Ltd.

Morley, L., McNally, A., Paszkiewicz, K., Corander, J., Méric, G., Sheppard, S.K., Blom, J. and Manning, G., 2015. Gene Loss and Lineage-Specific Restriction-Modification Systems Associated with Niche Differentiation in the Campylobacter jejuni Sequence Type 403 Clonal Complex. *Applied and Environmental Microbiology*, 81(11), pp.3641–3647.

Mourkas, E., Florez-Cuadrado, D., Pascoe, B., Calland, J.K., Bayliss, S.C.,

Mageiros, L., Méric, G., Hitchings, M.D., Quesada, A., Porrero, C., Ugarte-Ruiz, M., Gutiérrez-Fernández, J., Domínguez, L. and Sheppard, S.K., 2019. Gene pool transmission of multidrug resistance among Campylobacter from livestock, sewage and human disease. *Environmental Microbiology*, 00, pp.1462-2920.14760.

Murray, S., Pascoe, B., Méric, G., Mageiros, L., Yahara, K., Hitchings, M.D., Friedmann, Y., Wilkinson, T.S., Gormley, F.J., Mack, D., Bray, J.E., Lamble, S., Bowden, R., Jolley, K.A., Maiden, M.C.J., Wendlandt, S., Schwarz, S., Corander, J., Fitzgerald, J.R. and Sheppard, S.K., 2017. Recombination-mediated host adaptation by avian staphylococcus aureus. *Genome Biology and Evolution*, 9(4), pp.830–842.

Nachamkin, I., Yang, X.H. and Stern, N.J., 1993. Role of Campylobacter jejuni flagella as colonization factors for three-day-old chicks: analysis with flagellar mutants. *Applied and environmental microbiology*, 59(5), pp.1269–73.

Navarro-Gonzalez, N., Ugarte-Ruiz, M., Porrero, M.C., Zamora, L., Mentaberre, G., Serrano, E., Mateos, A., Lavín, S. and Domínguez, L., 2014. Campylobacter Shared Between Free-Ranging Cattle and Sympatric Wild Ungulates in a Natural Environment (NE Spain). *EcoHealth*, 11(3), pp.333–342.

Nothaft, H. and Szymanski, C.M., 2010. Protein glycosylation in bacteria: sweeter than ever. *Nature Reviews Microbiology*, 8(11), pp.765–778.

Novik, V., Hofreuter, D. and Galan, J.E., 2010. Identification of Campylobacter jejuni Genes Involved in Its Interaction with Epithelial Cells. *Infection and Immunity*, 78(8), pp.3540–3553.

Ochman, H., 2005. Genomes on the shrink. *Proceedings of the National Academy of Sciences*, 102(34), pp.11959–11960.

Ogden, I.D., Dallas, J.F., MacRae, M., Rotariu, O., Reay, K.W., Leitch, M., Thomson, A.P., Sheppard, S.K., Maiden, M., Forbes, K.J. and Strachan, N.J.C., 2009. Campylobacter Excreted into the Environment by Animal Sources: Prevalence, Concentration Shed, and Host Association. *Foodborne Pathogens and Disease*, 6(10), pp.1161–1170.

Pascoe, B., Williams, L.K., Calland, J.K., Meric, G., Hitchings, M.D., Dyer, M., Ryder, J., Shaw, S., Lopes, B.S., Chintoan-Uta, C., Allan, E., Vidal, A., Fearnley, C., Everest, P., Pachebat, J.A., Cogan, T.A., Stevens, M.P., Humphrey, T.J., Wilkinson, T.S., Cody, A.J., Colles, F.M., Jolley, K.A.,

Maiden, M.C.J., Strachan, N., Pearson, B.M., Linton, D., Wren, B.W., Parkhill, J., Kelly, D.J., van Vliet, A.H.M., Forbes, K.J. and Sheppard, S.K., 2019. Domestication of Campylobacter jejuni NCTC 11168. *Microbial genomics*, 5(7).

Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S.E. and Lercher, M.J., 2014. PopGenome: An Efficient Swiss Army Knife for Population Genomic Analyses in R. *Molecular Biology and Evolution*, 31(7), pp.1929–1936.

Price, M.N., Dehal, P.S. and Arkin, A.P., 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE*, 5(3), p.e9490.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.,

Rambaut, A., Drummond, A.J., Xie, D., Baele, G. and Suchard, M.A., 2018. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology*, 67(5), pp.901–904.

Rambaut, A., Lam, T.T., Max Carvalho, L. and Pybus, O.G., 2016. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evolution*, 2(1), p.vew007.

Santschi, D.E., Berthiaume, R., Matte, J.J., Mustafa, A.F. and Girard, C.L., 2005. Fate of Supplementary B-Vitamins in the Gastrointestinal Tract of Dairy Cows. *Journal of Dairy Science*, 88(6), pp.2043–2054.

Sheppard, S.K., Cheng, L., Méric, G., de Haan, C.P.A., Llarena, A.-K., Marttinen, P., Vidal, A., Ridley, A., Clifton-Hadley, F., Connor, T.R., Strachan, N.J.C., Forbes, K., Colles, F.M., Jolley, K.A., Bentley, S.D., Maiden, M.C.J., Hänninen, M.-L., Parkhill, J., Hanage, W.P. and Corander, J., 2014. Cryptic ecology among host generalist *Campylobacter jejuni* in domestic animals. *Molecular Ecology*, 23(10), pp.2442–2451.

Sheppard, S.K., Dallas, J.F., Strachan, N.J.C., MacRae, M., McCarthy, N.D., Wilson, D.J., Gormley, F.J., Falush, D., Ogden, I.D., Maiden, M.C.J. and Forbes, K.J., 2009. Campylobacter Genotyping to Determine the Source of Human Infection. *Clinical Infectious Diseases*, 48(8), pp.1072–1078.

Sheppard, S.K., Didelot, X., Meric, G., Torralbo, A., Jolley, K. a, Kelly, D.J., Bentley, S.D., Maiden, M.C.J., Parkhill, J. and Falush, D., 2013. Genome-wide association study identifies vitamin B5 biosynthesis as a host specificity factor in Campylobacter. *Proceedings of the National Academy of Sciences*, 110(29), pp.11923–11927.

Sheppard, S.K., Guttman, D.S. and Fitzgerald, J.R., 2018. Population genomics of bacterial host adaptation. *Nature Reviews Genetics*, 19(9), pp.549–565.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), pp.1312–1313.

Suttle, N.F., 1991. The Interactions Between Copper, Molybdenum, and Sulphur in Ruminant Nutrition. *Annual Review of Nutrition*, 11(1), pp.121–140.

Swanson, J., Kraus, S.J. and Gotschlich, E.C., 1971. Studies on gonococcus infection. I. Pili and zones of adhesion: their relation to gonococcal growth patterns. *The Journal of experimental medicine*, 134(4), pp.886–906.

Taboada, B., Ciria, R., Martinez-Guerrero, C.E. and Merino, E., 2012. ProOpDB: Prokaryotic Operon DataBase. *Nucleic Acids Research*, 40(D1), pp.D627–D631.

Tatusov, R.L., 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Research*, 28(1), pp.33–36.

Taylor, A.J., Zakai, S.A.I. and Kelly, D.J., 2017. The Periplasmic Chaperone Network of Campylobacter jejuni: Evidence that SalC (Cj1289) and PpiD (Cj0694) Are Involved in Maintaining Outer Membrane Integrity. *Frontiers in Microbiology*, 8(MAR), pp.1–14.

Thépault, A., Méric, G., Rivoal, K., Pascoe, B., Mageiros, L., Touzain, F., Rose, V., Béven, V., Chemaly, M. and Sheppard, S.K., 2017. Genome-Wide Identification of Host-Segregating Epidemiological Markers for Source Attribution in Campylobacter jejuni. *Applied and Environmental Microbiology*, 83(7), pp.e03085-16.

Vos, M. and Didelot, X., 2009. A comparison of homologous recombination rates in bacteria and archaea. *The ISME Journal*, 3(2), pp.199–208.

Wassenaar, T.M., van der Zeijst, B.A.M., Ayling, R. and Newell, D.G., 1993. Colonization of chicks by motility mutants of Campylobacter jejuni demonstrates the importance of flagellin A expression. *Journal of General Microbiology*, 139(6), pp.1171–1175.

Wilson, D.J., Gabriel, E., Leatherbarrow, A.J.H., Cheesbrough, J., Gee, S., Bolton, E., Fox, A., Hart, C.A., Diggle, P.J. and Fearnhead, P., 2009. Rapid evolution and the importance of recombination to the gastroenteric pathogen Campylobacter jejuni. *Molecular Biology and Evolution*, 26(2), pp.385–397.

Woodcock, D.J., Krusche, P., Strachan, N.J.C., Forbes, K.J., Cohan, F.M., Méric, G. and Sheppard, S.K., 2017. Genomic plasticity and rapid host switching can

promote the evolution of generalism: a case study in the zoonotic pathogen Campylobacter. *Scientific Reports*, 7(1), p.9650.

Woolhouse, M.E.J., Haydon, D.T. and Antia, R., 2005. Emerging pathogens: The epidemiology and evolution of species jumps. *Trends in Ecology and Evolution*, 20(5), pp.238–244.

Wu, S., Zhu, Z., Fu, L., Niu, B. and Li, W., 2011. WebMGA: a customizable web server for fast metagenomic sequence analysis. *BMC Genomics*, 12(1), p.444.

Data availability

Short read data for genomes sequenced as part of this chapter are archived on the NCBI SRA associated with BioProject accession PRJNA575343. Contiguous assemblies of all genome sequences compared are available at the public data repository FigShare (doi.org/10.6084/m9.figshare.9929054). Individual accession numbers can be found in the Appendix, Table S3.1.

Chapter 4

Gene pool transmission of multidrug resistance among *Campylobacter* from livestock, sewage and human disease

The work presented in this chapter is published as a peer-reviewed publication at:

Mourkas, E., Florez-Cuadrado, D., Pascoe, B., Calland, J.K., Bayliss, S.C., Mageiros, L., Méric, G., Hitchings, M.D., Quesada, A., Porrero, C., Ugarte-Ruiz, M., Gutierrez-Fernandez, J., Dominguez, L., Sheppard, S.K. 2019. "Gene pool transmission of multidrug resistance among *Campylobacter* from livestock, sewage and human disease." *Environmental Microbiology*. doi:10.1111/1462-2920.14760.

Commentary text

Work in chapter 1 has linked AMR as a major selective pressure on the evolution of *Campylobacter* species. The work in this chapter provides a detailed analysis of the AMR resistomes of *C. jejuni* and *C. coli* isolates from humans, livestock animals and urban effluents in Spain. Bacterial isolates were examined both *in vitro* and *in silico* for the presence of AMR. Work in this chapter identified the presence of genes conferring resistance to antibiotics and further characterized genetic associations of genes in MDR islands. These islands were often located on plasmids and mobile elements. Detailed allelic analysis showed AMR allele sharing between different sources and bacterial species with sewage playing a role in dissemination of resistance. Results in this chapter suggest the circulation of specific antibiotic resistance genes between humans, animals and the environment and provide further insights into the ecology of this major human bacterial pathogen. The statement of authorship for this chapter can be found in the Appendix supplementary form SF3.

Abstract

The use of antimicrobials in human and veterinary medicine has coincided with a rise in antimicrobial resistance (AMR) in the food-borne pathogens *Campylobacter jejuni* and *Campylobacter coli*. Faecal contamination from the main reservoir hosts (livestock, especially poultry) is the principal route of human infection but little is known about the spread of AMR among source and sink populations. In particular, questions remain about how *Campylobacter* resistomes interact between species and hosts, and the potential role of sewage as a conduit for the spread of AMR. Here the genomic variation associated with AMR in 168 *C. jejuni* and 92 *C. coli* strains isolated from humans, livestock and urban effluents in Spain, was investigated. Antimicrobial resistance was tested *in vitro* and isolate genomes were sequenced and screened for putative AMR genes and alleles. Genes associated with resistance to multiple drug classes were observed in both species and were commonly present in multidrug-resistant genomic islands (GIs), often located on plasmids or mobile elements. In many cases, these loci had alleles that were shared among *C. jejuni* and *C. coli* consistent with horizontal transfer. Results in this chapter suggest that specific antibiotic resistance genes have spread among *Campylobacter* isolated from humans, animals and the environment.

Introduction

Campylobacter is the leading cause of bacterial gastroenteritis in the European Union (EU) (Food and Authority, 2019). The most common pathogenic species, *C. jejuni* and *C. coli,* were responsible for over 245,658 cases of campylobacteriosis in the EU in 2016, surpassing disease caused by *E. coli, Salmonella* and *Listeria* (Food and Authority, 2019). *Campylobacter* are a common constituent of the gut microbiota of livestock including poultry, ruminants and pigs (Sheppard et al., 2009a; Sheppard et al., 2011; Sproston et al., 2011), and are also found in wild birds (Sheppard et al., 2010a; Sheppard et al., 2010b; Griekspoor et al., 2013; Cody et al., 2015; Atterby et al., 2018) and environmental sources (Dingle et al., 2001; Colles et al., 2003; Sheppard et al., 2009b). Human infection is typically associated with the consumption of contaminated meat (Fravalo et al., 2009; Hermans et al., 2012; Guyard-Nicodème et al., 2013) and causes acute gastroenteritis and is self-limiting after 3-5 days. In severe cases, antibiotic treatment may be required with fluoroquinolones and macrolides being the drugs of choice (Acheson and Allos, 2001).

Despite the ban on the use of antibiotics as growth promoters in animal feed in 2006 in the EU (Castanon, 2007), antimicrobial resistance (AMR) is still common among bacteria of the gastrointestinal tract of farmed animals (Sheppard et al., 2009a; Sheppard et al., 2009b; Sproston et al., 2011). According to the latest European Centre for Disease Prevention and Control (ECDC) report in 2017, *C. jejuni* and *C. coli* isolates of clinical and animal origin showed high levels of resistance to both ciprofloxacin and tetracycline (Food and Authority, 2019). Furthermore, *C. coli* from clinical and animal samples have displayed resistance to macrolides including erythromycin and the aminoglycoside streptomycin (Food and Authority, 2019). More worryingly, there is an apparent trend towards multidrug resistance (MDR), particularly among *C. coli* that regularly harbour different AMR genes simultaneously within the genome of a single isolate (Food and Authority, 2019; Pascoe et al., 2017; Luangtongkum et al., 2009).

Mechanisms of resistance are well documented for several drug classes including fluoroquinolones, tetracyclines, macrolides, aminoglycosides and β-lactams. Fluoroquinolone treatment was traditionally the first line of defence against campylobacteriosis but resistance has rapidly increased among strains (Sproston et al., 2018), potentially because it requires only a single point mutation in the genome (in the *gyrA* gene) (Gibreel, 2006; Payot et al., 2006; Luo et al., 2003; Luangtongkum et al., 2009). This has led to a shift in treatment in favour of erythromycin prescription (Nachamkin et al., 2000; Gibreel, 2006), where resistance arises from specific point mutations in 23S rRNA and develops relatively slowly (Lapierre et al., 2016). However, in 2014, erythromycin resistance was found in animal and clinical isolates that carried an rRNA methylating enzyme, the *ermB* gene (Qin et al., 2014; Wang et al., 2014). Two years later the *ermB* gene was detected in *C. coli* isolates from turkeys and chickens in Spain suggesting the mobilization of this gene through horizontal gene transfer (HGT) (Florez-Cuadrado et al., 2016; Florez-Cuadrado et al., 2018). Tetracycline resistance, associated with the *tetO* gene encoding a ribosomal protection protein, has also been observed in *Campylobacter* since 1987 (Sougakoff et al., 1987) and new enzymes conferring resistance to aminoglycosides continue to be discovered in *Campylobacter* (Iovine, 2013; Lambert et al., 1985; Zhao et al., 2016). In addition to these emerging trends, *Campylobacter* is known to have ‘natural’ resistance to β-lactams, such as penicillin,

in large part due to the ubiquity of the *bla*$_{\text{OXA-61}}$ gene (Alfredson and Korolik, 2005; Griggs et al., 2009). As a result of the widespread resistance to multiple antibiotic classes, it is no surprise that *Campylobacter* is a high priority pathogen on the recently published World Health Organization (WHO) list of bacteria, for which new antibiotics are urgently needed (WHO, 2017).

Many studies have highlighted the potential for transmission of AMR bacteria between agricultural animals and humans following extended use of antibiotics (Boerlin and Reid-Smith, 2008; Huttner et al., 2013). However, controversy surrounding evidence for a direct link is confounded by inconsistencies in interpreting what constitutes the spread of resistance. Broadly, the spread of AMR can be defined as a clonal transmission or gene pool transmission. In clonal transmission, bacteria that have acquired AMR in one niche are transmitted to another where they retain resistance, such as in the survival of resistant *Campylobacter* through the food production chain to infect humans (Yahara et al., 2017). In gene pool transmission, horizontal gene transfer (HGT) facilitates the spread of resistance genes between strains and species and the movement of genes (rather than clones) into multiple genetic backgrounds can be seen to spread AMR. Efforts to reduce AMR and conserve the remaining efficacy of existing drugs are focussed on the judicious use of antibiotics in animals and humans. In this context, it is advantageous to consider gene pool transmission as this is directly influenced by the selection pressure to maintain resistance in a given environment.

C. jejuni and *C. coli* can evolve rapidly, accumulating large numbers of nucleotide substitutions through mutation and recombination (Wilson et al., 2009; Sheppard et al., 2010a; Sheppard et al., 2010b; Dearlove et al., 2016). This can lead to *de novo* development of antimicrobial resistance through point mutation as well as the acquisition of resistance elements from other bacteria through HGT (Yahara et al., 2014; Yahara et al., 2016). HGT has a major role in the mobilization of AMR not only within bacterial species but even across species boundaries. For example, the *tetO* gene that confers resistance to tetracycline in *Campylobacter* (Batchelor, 2004; Taylor et al., 1983) is believed to have originated via HGT from a Gram-positive bacterium, potentially mediated by plasmid transfer (Taylor, 1986; Batchelor, 2004; Taylor et al., 1983). Interspecies genetic exchange requires some degree of niche overlap or physical proximity of strains. However, while there is some understanding

of host niche segregation and clonal transmission of particular *Campylobacter* lineages (Sheppard et al., 2009a; Sheppard et al., 2010a; Sheppard et al., 2010b; Sheppard et al., 2014), there is limited quantitative information about the transmission dynamics of AMR genes between human, animal and environmental gene pools (gene pool transmission) in this genus.

In this chapter, the genome of *Campylobacter* isolates from multiple sources from a survey of AMR in Spain, was sequenced, to consider the following questions: (1) What is the MDR profile of *C. jejuni* and *C. coli* isolates? (2) How is AMR distributed across *Campylobacter* source and sink populations? (3) Are there species or lineages that harbour more AMR? (4) Is there movement of AMR between animal, human and environmental gene pools? MDR resistance phenotypes are quantified *in vitro* and compared to putative genomic determinants identified from over 2,000 known AMR genes. The co-localization of these genes within resistance islands is examined and the allelic variation is compared among isolates from different sample sources. These analyses provide a basis for considering the interaction of different AMR gene pools and the potential source/sink contribution of livestock, humans and sewage effluents to the *Campylobacter* resistome.

Results

Enhanced *in vitro* multidrug resistance in *C. coli* compared to *C. jejuni*

A collection of 168 *C. jejuni* and 92 *C. coli* isolates of human, animal and sewage origin, was assembled (Appendix, Table S4.1). *In vitro* resistance to six antibiotics (ciprofloxacin, nalidixic acid, tetracycline, erythromycin, streptomycin and gentamicin) of isolates of animal origin (Table 4.1, Appendix, Table S4.2) was compared to resistance profiles of isolates of human and sewage origin (Table 4.1, Appendix, Table S4.2). All *Campylobacter* isolates that were resistant to both ciprofloxacin and nalidixic acid were referred to as ciprofloxacin resistant only because resistance is conferred by SNPs in the same gene. The highest proportion of antimicrobial resistance was to ciprofloxacin (146/163; 90.1% for *C. jejuni* and 86/91; 94.5% for *C. coli*) and tetracycline (149/163; 91.4% for *C. jejuni* and 86/91; 94.5% for *C. coli*), followed by streptomycin (24/163; 14.7% for *C. jejuni* and 58/91; 63.7% for *C. coli*), erythromycin (4/162; 2.5% for *C. jejuni* and 23/91; 25.3% for *C. coli*) and gentamicin (2/163; 1.2% for *C. jejuni* and 10/91; 11% for *C. coli*) (Table 4.1, Appendix, Table S4.2).

Table 4.1. Drug resistance profiles of 254 *Campylobacter* isolates from humans, animals and sewage.

	C. jejuni				*C. coli*			
Antibiotics[a]	Animals	Humans	Sewage	Total	Animals	Humans	Sewage	Total
Ciprofloxacin	36/44 (81.8%)	106/115 (88.7%)	4/4 (100%)	146/163 (90.12%)	11/11 (100%)	32/33 (97%)	43/47 (91.5%)	86/91 (94.5%)
Nalidixic acid	35/44 (79.54%)	78/115 (67.83%)	3/4 (75%)	116/163 (71.16%)	11/11 (100%)	30/33 (90.1%)	43/47 (91.5%)	84/91 (92.31%)
Tetracycline	39/44 (88.6%)	108/115 (93.91%)	2/4 (50%)	149/163 (91.41%)	11/11 (100%)	31/33 (94%)	44/47 (93.6%)	86/91 (94.5%)
Erythromycin	3/44 (6.8%)	1/115 (0.87%)	0/4 (0%)	4/163 (2.45%)	10/11 (90.1%)	6/33 (18.2%)	7/47 (14.9%)	23/91 (25.3%)
Streptomycin	15/44 (34.1%)	9/115 (7.83%)	0/4 (0%)	24/163 (14.72%)	10/11 (90.1%)	18/33 (54.5%)	30/47 (63.8%)	58/91 (63.7%)
Gentamicin	0/44 (0%)	2/115 (1.7%)	0/4 (0%)	2/163 (1.23%)	4/11 (36.4%)	2/33 (6.1%)	4/47 (8.51%)	10/91 (11%)
Total number of isolates	44	115	4	163	11	33	47	91

[a]Antibiotics: C, ciprofloxacin; T, tetracycline; E, Erythromycin; S, streptomycin; G, gentamicin

Higher prevalence of resistance was observed in *C. coli* isolates to erythromycin, streptomycin and gentamicin compared to *C. jejuni* (Fisher's exact test; $p<0.001$). Typically, an isolate is considered multidrug resistant when it is resistant to at least three different classes of antibiotics (European Centre for Disease Prevention and Control [ECDC] & European Food Safety Authority [EFSA], 2015). Based on this criterion, more *C. coli* isolates were MDR (49/91; 53.8%) than *C. jejuni* (27/163; 16.6%) (Table 4.2). All *C. coli* isolates were resistant to at least one antibiotic (Table 4.2). Six (out of 163; 3.7%) *C. jejuni* isolates were sensitive to all tested antibiotics. Most of the isolates tested were resistant to both ciprofloxacin and tetracycline (140/163 or 85.9% *C. jejuni* and 82/91 or 90.1% *C. coli*), of which 52 *C. coli* isolates (57.1%) were also resistant to streptomycin compared to 24 *C. jejuni* isolates (14.7%) and 9 *C. coli* isolates (9.9%) were also resistant to gentamicin compared to 2 *C. jejuni* isolates (1.23%) (Table 4.2). [N.B. The above section of data analysis was performed by Diego Flórez Cuadrado by the Universidad Complutense Madrid, Madrid, Spain.]

AMR isolates are distributed across highly structured populations

High levels of AMR observed in laboratory assays could indicate either an abundance of low diversity AMR clones or proliferation of AMR in multiple lineages. To investigate this, the population genomic structure of AMR isolates was analysed. The core genome phylogeny revealed that AMR isolates belonged to genome sequence clusters consistent with existing multilocus sequence typing (MLST) sequence type (ST) and clonal complex designations (Dingle et al., 2001; Miller, 2006) (Figure 4.1). *C. jejuni* isolates of chicken and cattle origin were mainly of host generalist (ST-21, ST-48, ST-206 and ST-45) clonal complexes (Sheppard et al., 2010a; Sheppard et al., 2014) (Figure 4.1, Appendix, Table S4.1). Cattle isolates also belonged to ST-61 and ST-42 cattle associated clonal complexes, while human clinical isolates contained isolates of these generalist and cattle associated clonal complexes as well as additional generalist clonal complexes (ST-22, ST-52) and chicken associated clonal complexes (ST-257, ST-353, ST-354, ST-443, ST464, ST-574 and ST-658) (Figure 4.1, Appendix, Table S4.1). *C. jejuni* isolates from sewage belonged to ST-362, a human associated complex and generalist ST-22, ST-45 and ST-607 complexes (Figure 4.1, Appendix, Table S4.1). Multidrug resistant *C. jejuni* isolates (27/167) were from generalist (ST-21, ST-206, ST-45, ST-52) complexes, chicken associated complexes (ST-354, ST-460 and ST-464)

and cattle associated complexes (ST-42 and ST-61) (Figure 4.1, Appendix, Table S4.1). *C. coli* isolates represented 28 different STs, all of which belonged to the ST-828 clonal complex. The most abundant STs were 825 and 827, constituting 20.7% and 17.4% of all *C. coli* isolates (Figure 4.2, Appendix, Table S4.1). The proportion of *C. coli* isolates displaying multidrug resistance (60.9%) was considerably higher than within *C. jejuni* (16.1%), nearly half of which were isolated from sewage highlighting the potential importance of urban effluents as reservoirs of AMR genes (Figure 4.2, Table 4.2). Clearly, diversity within this complex is lower than in agricultural/clinical *C. jejuni* and one might consider ST-828 complex to be a single clone. However, as illustrated (Figure 4.2), AMR is found in divergent lineages within the ST-828 complex and, importantly, is also absent in some closely related strains. This pattern is inconsistent with the proliferation of a clone that acquired AMR genes in a single ancestral acquisition event. Rather it suggests horizontal transfer of AMR genes among sublineages. [N.B. The above section of data analysis was performed in collaboration with Diego Flórez Cuadrado by the Universidad Complutense Madrid, Madrid, Spain.]

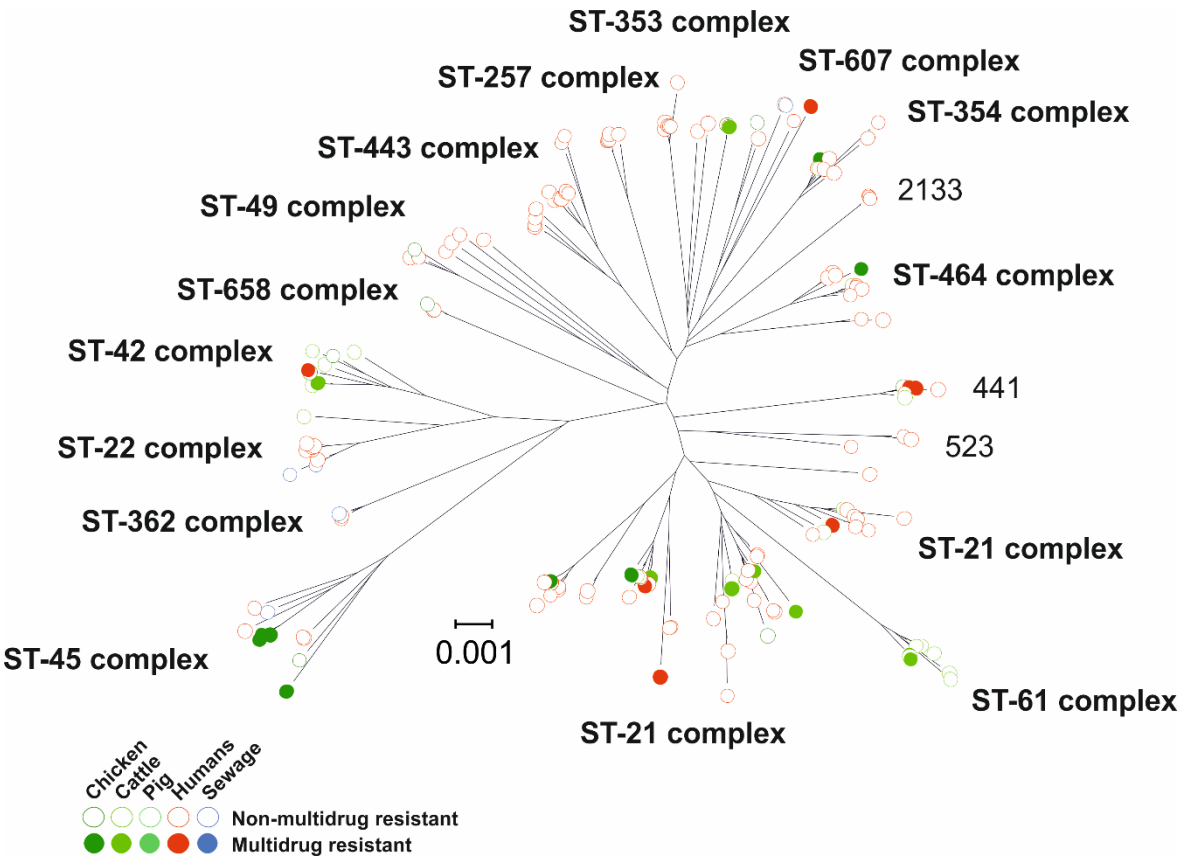


Figure 4.1. Phylogeny of antimicrobial resistant *Campylobacter jejuni*. Tree was reconstructed for 167 *C. jejuni* using concatenated gene-by-gene alignments of 595 core genes using the neighbour-joining algorithm. Common sequence types (STs) and clonal complexes, defined by

MLST, are indicated on the trees. Multidrug resistant isolates from chickens (dark green), cattle (intermediate green), pigs (light green), humans (red) and sewage (blue) are indicated with a filled circle, while the non-multidrug resistant isolates are indicated with an open circle. The scale bar represents the number of substitutions per site.

Table 4.2. Multidrug resistant and non-multidrug resistant *Campylobacter* isolates (n=254) from humans, animals and sewage.

		***C. jejuni* (n=162)**			***C. coli* (n=91)**		
	Antibiotics[a]	Animals	Humans	Sewage	Animals	Humans	Sewage
Multi resistant	**CTESG**	-	-	-	4/11 (36.4%)	1/33 (3%)	-
	CTES	-	-	-	5/11 (45.5%)	4/33 (12.1%)	5/47 (10.6)
	CTSG	-	2/115 (1.7%)	-	-	1/33 (3%)	3/47 (6.4%)
	CTS	15/44 (34.1%)	7/115 (6.9%)	-	1/11 (9.1%)	11/33 (33.3%)	17/47 (36.2%)
	CTE	2/44 (4.5%)	1/115 (0/9%)	-	1/11 (9.1%)	1/33 (3%)	2/47 (4.3%)
	CT	16/44 (36.4%)	95/115 (82.6%)	2/4 (50%)	-	12/33 (36.4%)	13/47 (27.7%)
Non-multi resistant	CS	-	-	-	-	-	1/47 (2.1%)
	TE	1/44 (2.27%)	-	-	-	-	-
	TS	1/44 (2.27%)	-	-	-	1/33 (3%)	4/47 (8.5%)
	C	3/44 (6.8%)	1/115 (0.9%)	1/4 (25%)	-	2/33 (6.1%)	2/47 (4.25%)
	T	4/44 (11.4%)	5/115 (4.4 %)	-	-	-	-
Non resistant	Sensitive	2/44 (4.5%)	4/115 (3.5%)	-	-	-	-
Total number of non-multidrug resistant		27/44 (61.36%)	101/115 (8.69%)	4/4 (100 %)	-	15/33 (45.45%)	27/47 (57.44%)
Total number of multidrug resistant		17/44 (38.63%)	10/115 (87.82%)	-	11/11 (100%)	18/33 (54.54%)	20/47 (42.55%)
Total number of isolates		44	115	4	11	33	47

[a]Antibiotics: C, ciprofloxacin; T, tetracycline; E, erythromycin; S, streptomycin; G, gentamicin

C. coli genomes harbour more antimicrobial resistance genes than *C. jejuni*

The genome sequences of all *Campylobacter* isolates were compared to 2,158, 2,280 and 4,324 known antibiotic resistance genes and alleles from the Comprehensive Antibiotic Resistance Database (CARD) (Cameron and Gaynor, 2014), ResFinder (Zankari et al., 2012) and the National Center for Biotechnology Information (NCBI) databases, respectively. The analysis revealed the presence of 18 AMR genes including: *cmeA, cmeB, cmeC,* bla_{OXA-61}, *tetO, ant-like* A*, ant-like* B*,*

ant(6)-Ia, sat-1, sat-4, lnuC, ant(6)-Ib, aad9, aph(3)-IIIa, aph(2)-IIIa, hpt, apmA and *ermB* (Figure 4.3, Figure 4.4, Table 4.3) (Qin et al., 2012; Florez-Cuadrado et al., 2016; Toth et al., 2013; Yao et al., 2017; Olkkola et al., 2016; Achard et al., 2005; Alfredson and Korolik, 2005; Trieu-Cuot et al., 1985; Griggs et al., 2009; Zhao et al., 2016; Sougakoff et al., 1987; Cameron and Gaynor, 2014).

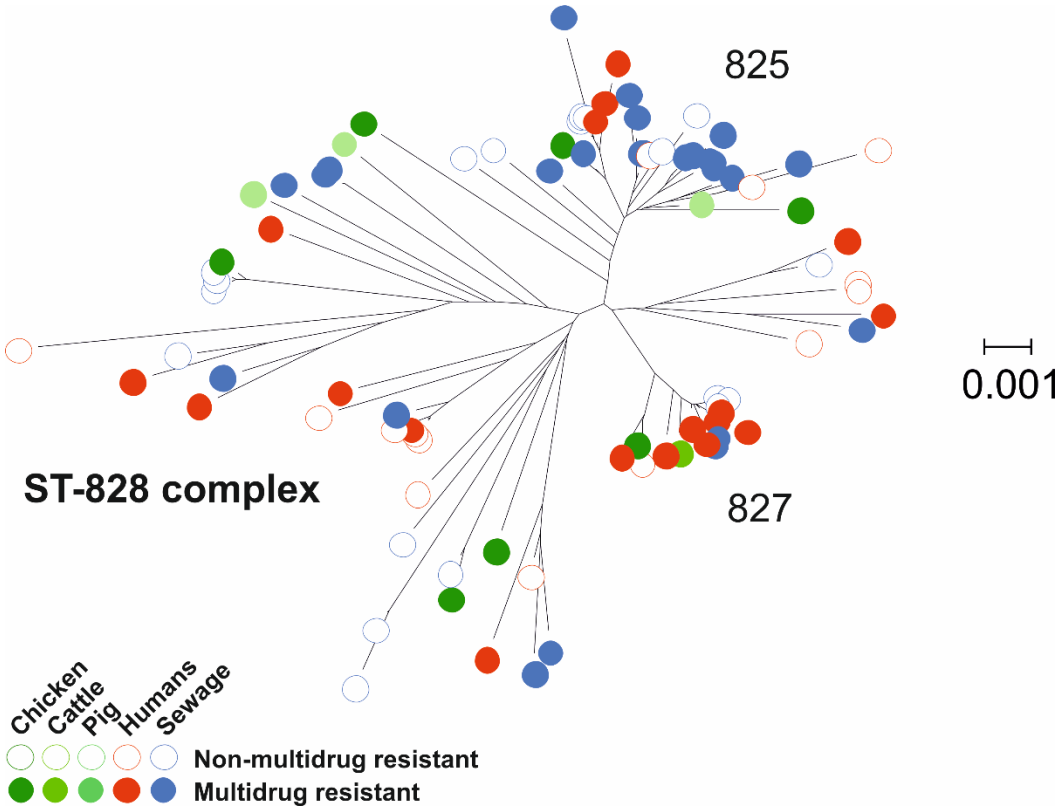


Figure 4.2. Phylogeny of antimicrobial resistant *Campylobacter coli.* Tree was reconstructed for 92 *C. coli* using concatenated gene-by-gene alignments of 595 core genes using the neighbour-joining algorithm. Common sequence types (STs) and clonal complexes, defined by MLST, are indicated on the trees. Multidrug resistant isolates from chickens (dark green), cattle (intermediate green), pigs (light green), humans (red) and sewage (blue) are indicated with a filled circle, while the non-multidrug resistant isolates are indicated with an open circle. The scale bar represents the number of substitutions per site.

The *cmeA*, *cmeB* and *cmeC* genes, associated with efflux pump function, were present in all isolates. The bla_{OXA-61} and *tetO* genes were common in resistant *C. jejuni* and *C. coli* isolates (Figure 4.3, Figure 4.4, Table 4.3). The genes *ant-like* A and *ant-like* B have been described before as separate genes (Olkkola et al., 2016) and later revised as *ant(6)-Ie* (Hormeño et al., 2018). To avoid the issues of gene duplication and gene paralogues they are considered as separate genes in this chapter. The bla_{OXA-61} gene was significantly more prevalent in *C. jejuni* (64.8%) than *C. coli* isolates (51.1%) (Fisher's exact test; $p<0.05$), while the *ant-like* A gene was more prevalent in *C. coli* (40.22% of *C. coli* and 1.19% of *C. jejuni* isolates,

p<0.001). The prevalence of the *ant-like* A gene was also significantly higher in multidrug resistant isolates (33.7%) compared to non-multidrug resistant isolates (6.7%) (p<0.001) (Figure 4.3, Figure 4.4, Table 4.3), and associated (p<0.005) with isolates from humans (14.5%) and sewage (13.3%) compared to those from animals (1.2%) (p<0.005) (Figure 4.3, Figure 4.4, Table 4.3).

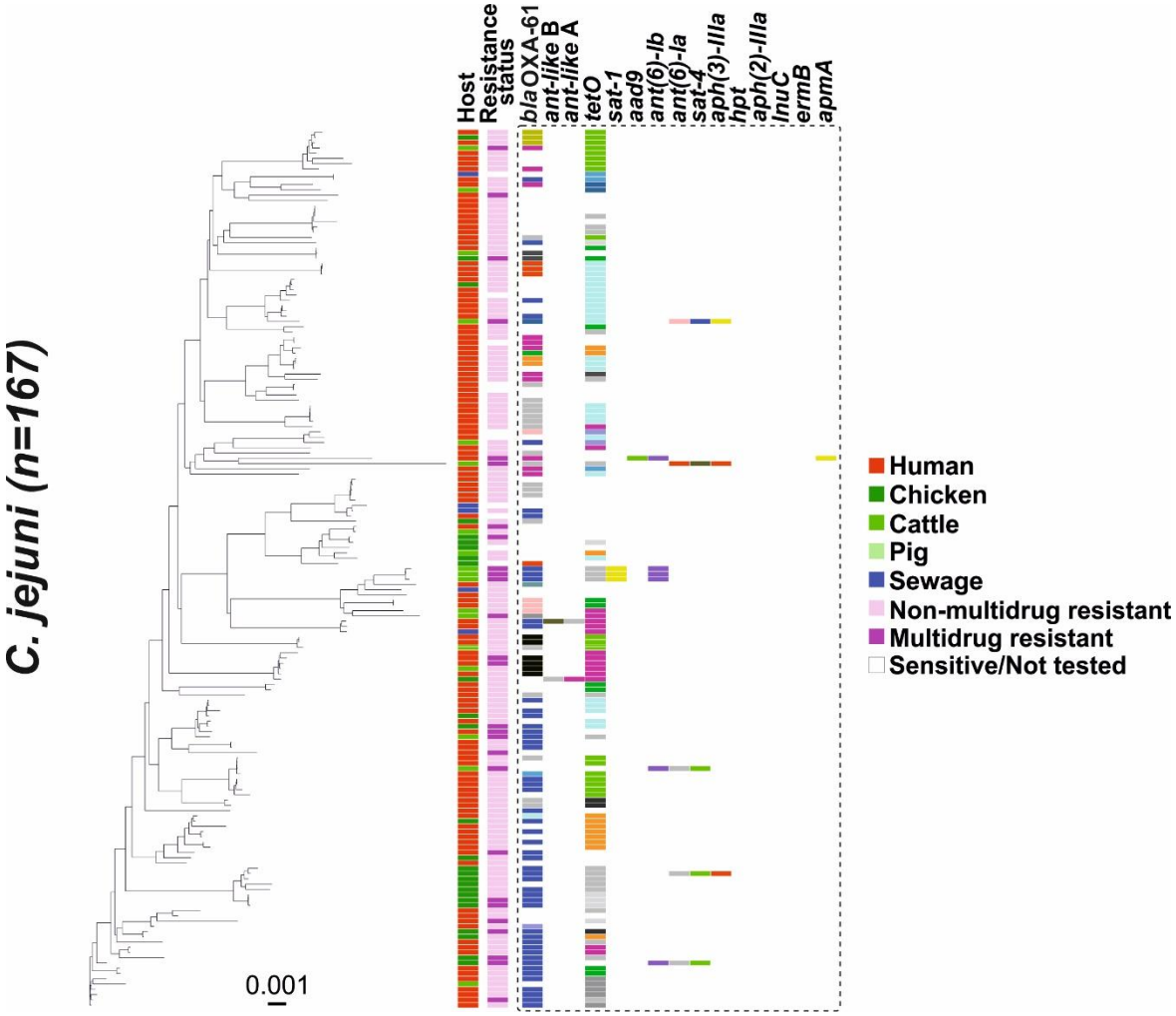


Figure 4.3. Presence and allelic diversity of 15 antimicrobial resistance genes in *C. jejuni* genomes. Phylogenetic tree was reconstructed using gene-by-gene concatenated alignments of 595 core genes, and the neighbour-joining algorithm for 167 *C. jejuni*. Isolate source is shown in the first column for chicken (dark green), cattle (green), pigs (light green), humans (red) and sewage (blue). The second column indicates the resistance status of each isolate as multidrug resistant (dark pink), non-multidrug resistance (light pink) or not tested (white). Remaining columns indicate allelic variation at known resistance gene loci, with identical alleles coloured with the same colour. The scale represents the number of substitutions per site.

In the case of non-multidrug resistant isolates, the frequency difference of the *ant-like* A gene can probably be attributed to the frequency of *C. jejuni* in human infection samples compared to the abundance of *C. coli* from sewage. Genes associated with aminoglycoside resistance (*ant(6)-Ia, sat-4, ant(6)-Ib, aad9, aph(3)-IIIa, aph(2)-IIIa,*

hpt and *apmA*) were mainly found in *C. coli* multidrug resistant isolates while *sat-1* was detected in only 3 *C. jejuni* strains from animals (Figure 4.3, Figure 4.4, Table 4.3). Genes *ant(6)-Ia, sat-4, ant(6)-Ib* and *aph(3)-IIIa* were also found in *C. jejuni* isolates from animals (Figure 4.2, Table 4.3). The *lnuC* gene, conferring resistance to lincosamides, was detected only in *C. coli* isolates and the *ermB* gene, which is not commonly found in *Campylobacter*, was detected in only one *C. coli* isolate from a chicken (Figure 4.4, Table 4.3). A strong positive correlation ($p<0.001$) between resistance phenotypes and genotypes was observed for tetracycline, streptomycin and gentamicin that were tested *in vitro* (Appendix, Table S4.3). There was no correlation for erythromycin because the associated AMR gene *ermB* was only found in one isolate (Appendix, Table S4.3).

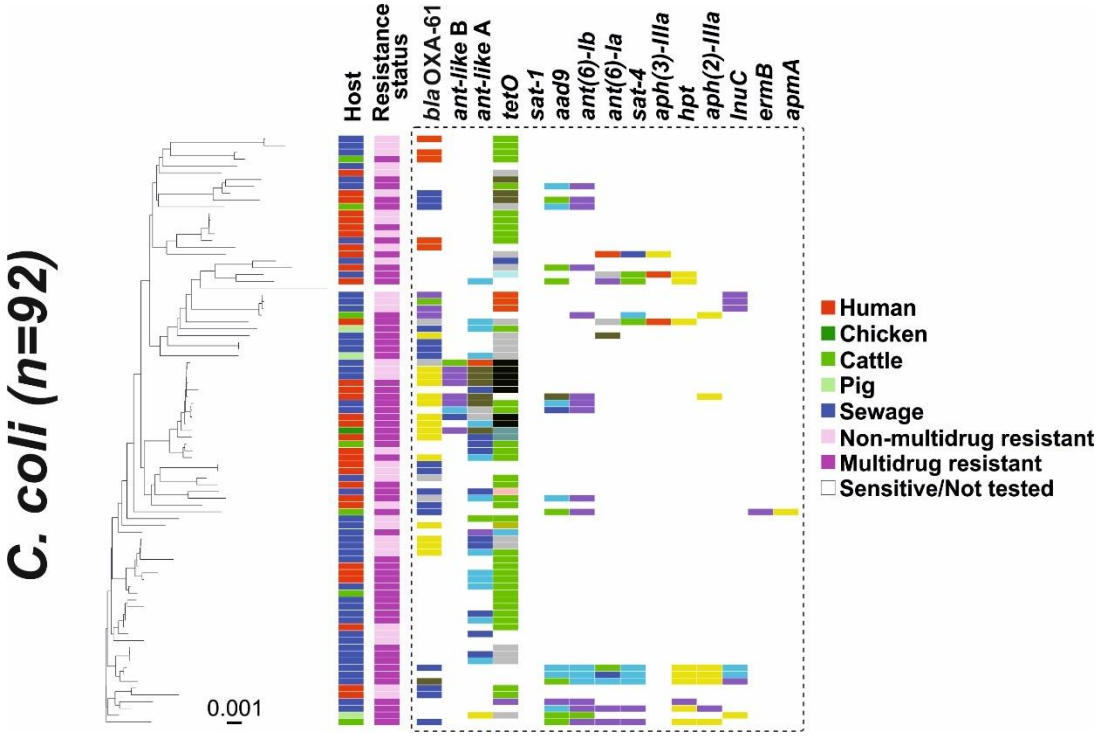


Figure 4.4. Presence and allelic diversity of 15 antimicrobial resistance genes in *C. coli* genomes. Phylogenetic tree was reconstructed using gene-by-gene concatenated alignments of 595 core genes, and the neighbour-joining algorithm for 92 *C. coli*. Isolate source is shown in the first column for chicken (dark green), cattle (green), pigs (light green), humans (red) and sewage (blue). The second column indicates the resistance status of each isolate as multidrug resistant (dark pink), non-multidrug resistance (light pink) or not tested (white). Remaining columns indicate allelic variation at known resistance gene loci, with identical alleles coloured with the same colour. The scale represents the number of substitutions per site.

Concordance between putative resistance genotypes and laboratory phenotypes was lower than in some previous studies (Zhao et al., 2016; Tyson et al., 2015; McDermott et al., 2016). The main reason for this was that this chapter principally

focused on the differential presence of AMR genes, to understand gene pool transmission, rather than resistance conferred by point mutation where it is more difficult to differentiate horizontal acquisition from *de novo* mutation. Other incongruences were observed between genotype prediction and laboratory phenotype. For example, not all isolates carrying aminoglycoside resistance genes were phenotypically resistant to streptomycin and gentamicin (Appendix, Table S4.3). This is consistent with previous studies (Tyson et al., 2015; McDermott et al., 2016) and is potentially associated with variation in gene expression levels or synergistic effects among different resistance genes, warranting further study. [N.B. The above section of data analysis was performed in collaboration with Diego Flórez Cuadrado by the Universidad Complutense Madrid, Madrid, Spain.]

AMR genes are co-localized in the genome of multidrug resistant isolates

AMR genes are often found in close proximity in the genome. For example, aminoglycoside resistance genes can form localized clusters within the genome (Werner et al., 2003; Qin et al., 2012). The low numbers of *apmA* and *ermB* genes identified, excluded them from formal statistical comparison. Due to the high levels of resistance to fluoroquinolones and tetracycline, the presence of *ant-like* A*, ant(6)-Ia, sat-4, ant(6)-Ib, aad9, aph(3)-IIIa, aph(2)-IIIa, sat-1* and *hpt* genes, was by definition significantly associated with multidrug resistance (Fisher's exact test; $p<0.001$), because this was defined as resistance to three or more antimicrobial classes (Table 4.3). There was a slight increasing trend in the presence of *ant-like* A*, ant-like* B*, aad9, ant(6)-Ia, sat-4, ant(6)-Ib,* and *aph(3)-IIIa* genes from 2010 to 2015 (Appendix, Figure S4.2). Furthermore, the relative position of the 15 AMR genes (in contiguous sequence assemblies) detected in *Campylobacter* isolates revealed two types of genetic associations in animal, human and sewage isolates. The first was between *ant(6)-Ia, sat-4* and *aph(3)-IIIa* genes, which clustered together in three *C. jejuni* isolates (1 from chicken and 2 from cattle) and in eight *C. coli* isolates (1 from chicken, 4 from humans and 3 from sewage) (Figure 4.5). This cluster has been previously described with the three genes located on the same genomic island in *C. coli* (Qin et al., 2012). The further addition of the *aph(2)-IIIa* gene to this genomic island was observed in two *C. coli* isolates from sewage (Figure 4.5). The second type of genetic association involved the presence of *tetO, aad9* and *ant(6)-Ib* genes. These genes clustered together in six *C. coli* isolates (1 from chicken, 1 from pig, 1 from human and 3 from sewage) but also in one *C. jejuni*

Table 4.3. Prevalence of 15 AMR genes in 168 *C. jejuni* and 92 *C. coli* isolates.

	Multidrug resistant					Non-multidrug resistant					Sensitive			Not tested		
	C. jejuni (n=27)		*C. coli* (n=56)			*C. jejuni* (n=129)			*C. coli* (n=35)		*C. jejuni* (n=7)			*C. jejuni* (n=5)**		*C. coli* (n=1)**
	Animals (n=17)	Humans (n=10)	Animals (n=11)	Humans (n=18)	Sewage (n=27)	Animals (n=25)	Humans (n=101)	Sewage (n=3)	Humans (n=15)	Sewage (n=20)	Animals (n=2)	Humans (n=4)	Sewage (n=1)	Humans (n=3)	Sewage (n=2)	Humans (n=1)
*bla*OXA-61	15/17 (88.24%)	6/10 (60.00%)	8/11 (72.73%)	9/18 (50.00%)	9/27 (33.33%)	16/25 (64.00%)	65/101 (64.36%)	1/3 (33.33%)	8/15 (53.33%)	13/20 (65.00%)	1/2 (50.00%)	3/4 (75.00%)	0/1 (0%)	1/3 (33.33%)	0/2 (0.00%)	0/1 (0.00%)
*tet*O	14/17 (82.35%)	4/10 (40.00%)	8/11 (72.73%)	16/18 (88.89%)	23/27 (85.19%)	20/25 (80.00%)	80/101 (79.21%)	1/3 (33.33%)	12/15 (80.00%)	16/20 (80.00%)	0/2 (0.00%)	0/4 (0/00%)	0/1 (0%)	3/3 (100.00%)	1/2 (50.00%)	1/1 (100.00%)
ant-like B	0/17 (0.00%)	0/10 (0.00%)	1/11 (9.09%)	3/18 (16.67%)	2/27 (7.41%)	1/25 (4.00%)	1/101 (0.99%)	0/3 (0.00%)	0/15 (0.00%)	3/20 (15.00%)	0/2 (0.00%)	0/4 (0/00%)	0/1 (0%)	0/3 (0.00%)	0/2 (0.00%)	0/1 (0.00%)
ant-like A	0/17 (0.00%)	0/10 (0.00%)	5/11 (45.45%)	12/18 (66.67%)	11/27 (40.74%)	1/25 (4.00%)	1/101 (0.99%)	0/3 (0.00%)	1/15 (6.67%)	8/20 (40.00%)	0/2 (0.00%)	0/4 (0/00%)	0/1 (0%)	0/3 (0.00%)	0/2 (0.00%)	0/1 (0.00%)
ant(6)-Ia	4/17 (23.53%)	0/10 (0.00%)	1/11 (9.09%)	3/18 (16.67%)	6/27 (22.22%)	1/25 (4.00%)	0/101 (0.00%)	0/3 (0.00%)	1/15 (6.67%)	0/20 (0.00%)	0/2 (0.00%)	0/4 (0/00%)	0/1 (0%)	0/3 (0.00%)	0/2 (0.00%)	1/1 (100.00%)
sat-4	4/17 (23.53%)	0/10 (0.00%)	2/11 (18.18%)	3/18 (16.67%)	5/27 (18.52%)	1/25 (4.00%)	0/101 (0.00%)	0/3 (0.00%)	1/15 (6.67%)	0/20 (0.00%)	0/2 (0.00%)	0/4 (0/00%)	0/1 (0%)	0/3 (0.00%)	0/2 (0.00%)	1/1 (100.00%)
lnuC	0/17 (0.00%)	0/10 (0.00%)	1/11 (9.09%)	0/18 (0.00%)	3/27 (11.11%)	0/25 (0.00%)	0/101 (0.00%)	0/3 (0.00%)	1/15 (6.67%)	4/20 (20.00%)	0/2 (0.00%)	0/4 (0/00%)	0/1 (0%)	0/3 (0.00%)	0/2 (0.00%)	0/1 (0.00%)
ant(6)-Ib	5/17 (29.41%)	1/10 (10.00%)	5/11 (45.45%)	4/18 (22.22%)	8/27 (29.63%)	0/25 (0.00%)	0/101 (0.00%)	0/3 (0.00%)	0/15 (0.00%)	0/20 (0.00%)	0/2 (0.00%)	0/4 (0/00%)	0/1 (0%)	0/3 (0.00%)	0/2 (0.00%)	0/1 (0.00%)
aad9	0/17 (0.00%)	1/10 (10.00%)	4/11 (36.36%)	5/18 (27.78%)	8/27 (29.63%)	0/25 (0.00%)	0/101 (0.00%)	0/3 (0.00%)	0/15 (0.00%)	0/20 (0.00%)	0/2 (0.00%)	0/4 (0/00%)	0/1 (0%)	0/3 (0.00%)	0/2 (0.00%)	0/1 (0.00%)
aph(3)-IIIa	2/17 (11.76%)	0/10 (0.00%)	0/11 (0.00%)	2/18 (11.11%)	1/27 (3.7%)	1/25 (4.00%)	0/101 (0.00%)	0/3 (0.00%)	1/15 (6.67%)	0/20 (0.00%)	0/2 (0.00%)	0/4 (0/00%)	0/1 (0%)	0/3 (0.00%)	0/2 (0.00%)	1/1 (100.00%)
aph(2)-IIIa	0/17 (0.00%)	0/10 (0.00%)	2/11 (18.18%)	1/18 (5.56%)	4/27 (14.81%)	0/25 (0.00%)	0/101 (0.00%)	0/3 (0.00%)	0/15 (0.00%)	0/20 (0.00%)	0/2 (0.00%)	0/4 (0/00%)	0/1 (0%)	0/3 (0.00%)	0/2 (0.00%)	0/1 (0.00%)
hpt	0/17 (0.00%)	0/10 (0.00%)	1/11 (9.09%)	2/18 (11.11%)	6/27 (22.22%)	0/25 (0.00%)	0/101 (0.00%)	0/3 (0.00%)	0/15 (0.00%)	0/20 (0.00%)	0/2 (0.00%)	0/4 (0/00%)	0/1 (0%)	0/3 (0.00%)	0/2 (0.00%)	0/1 (0.00%)
apmA	0/17 (0.00%)	1/10 (10.00%)	1/11 (9.09%)	0/18 (0.00%)	0/27 (0.00%)	0/25 (0.00%)	0/101 (0.00%)	0/3 (0.00%)	0/15 (0.00%)	0/20 (0.00%)	0/2 (0.00%)	0/4 (0/00%)	0/1 (0%)	0/3 (0.00%)	0/2 (0.00%)	0/1 (0.00%)
sat-1	3/17 (17.65%)	0/10 (0.00%)	0/11 (0.00%)	0/18 (0.00%)	0/27 (0.00%)	0/25 (0.00%)	0/101 (0.00%)	0/3 (0.00%)	0/15 (0.00%)	0/20 (0.00%)	0/2 (0.00%)	0/4 (0/00%)	0/1 (0%)	0/3 (0.00%)	0/2 (0.00%)	0/1 (0.00%)
ermB	0/17 (0.00%)	0/10 (0.00%)	1/11 (9.09%)	0/18 (0.00%)	0/27 (0.00%)	0/25 (0.00%)	0/101 (0.00%)	0/3 (0.00%)	0/15 (0.00%)	0/20 (0.00%)	0/2 (0.00%)	0/4 (0/00%)	0/1 (0%)	0/3 (0.00%)	0/2 (0.00%)	0/1 (0.00%)

*Isolates are separated as multidrug or non-multidrug resistant based on their *in vitro* phenotypic profile

**Isolates id: 5087, 5093, 5111, 5095, 5100, 5215 were not tested for antibiotic resistant profile *in vitro*

isolate from a human patient (Figure 4.5). The addition of the *sat-1, hpt, apmA* and *ermB* genes was also observed in these two types of syntenic block (Figure 4.5). [N.B. The above section of data analysis was performed in collaboration with Diego Flórez Cuadrado by the Universidad Complutense Madrid, Madrid, Spain.]

Evidence of gene pool transmission AMR genes

Evidence for horizontal gene transfer has been demonstrated for AMR genes in various bacteria, including *Campylobacter* (Wang et al., 2014; Sheppard et al., 2013; Sheppard and Maiden, 2015; Sheppard et al., 2011; Li et al., 2017), in some cases facilitated by mobile genetic elements including plasmids and transposons (Boerlin and Reid-Smith, 2008). One plasmid (pCFSAN032805; Accession: CP023546.1) was identified in the genome sequences of 8 *C. coli* isolates (1 from chicken, 1 from a pig, 3 from humans and 3 from sewage) (Figure 4.5). Furthermore, a *C. jejuni* plasmid (pCJ14980A; Accession: CP017030.1) previously isolated from turkey faeces (Florez-Cuadrado et al., 2017) was identified in a *C. jejuni* isolate from cattle in this chapter (Figure 4.5). A pTet plasmid (Accession: CP002030.1) was also detected in one *C. coli* isolate of human origin (Figure 4.5). A genomic region that was carrying the gene cluster *ant(6)-Ia, sat-4* and *aph(3)-IIIa* was highly similar to an integrative conjugative mobile element described in *Erysipelothrix rhusiopathiae* (Accession: MG812141.1) isolated from a pig farm. This region was also similar to sequences from other bacteria like *Clostridium difficile, Staphylococcus aureus, Staphylococcus pseudintermedius, Streptococcus suis* and *Enterococcus faecium.* These findings are consistent with the circulation of genes, and more specifically alleles, not only between host microbiome gene pools but also between *Campylobacter* species. To investigate this further, the allelic diversity for the 15 identified AMR genes in *C. jejuni* and *C. coli* isolates, was compared.

The genes, *bla*$_{\text{OXA-61}}$ and *tetO,* had the highest diversity with 34 and 47 different alleles detected in *C. jejuni* and in *C. coli* isolates respectively (Figure 4.5, Appendix, Figure S4.1). There were five *bla*$_{\text{OXA-61}}$ alleles, two of which were present in 16 and four *C. jejuni* and in 50 and five *C. coli* isolates, respectively (Figure 4.5, Appendix, Figure S4.1). For the *tetO* gene, six alleles were present in more than five isolates each, with the most common allele present in 19 *C. jejuni* and in 35 *C. coli*. For the *aad9* and *ant(6)-Ib* gene, both of which had five alleles, the most common allele was present in both *C. jejuni* and *C. coli* isolates from multiple sources (Figure 4.5,

Appendix, Figure S4.1, Table S4.2). Finally, the *sat-4* gene shared two out of the six alleles between four *C. jejuni* and four *C. coli* isolates and the *apmA* gene had one allele which was shared by a *C. jejuni* of human origin and a *C. coli* isolated from a chicken (Figure 4.5, Appendix, Figure S4.1, Table S4.2). Remaining alleles were detected exclusively in *C. coli* isolates.

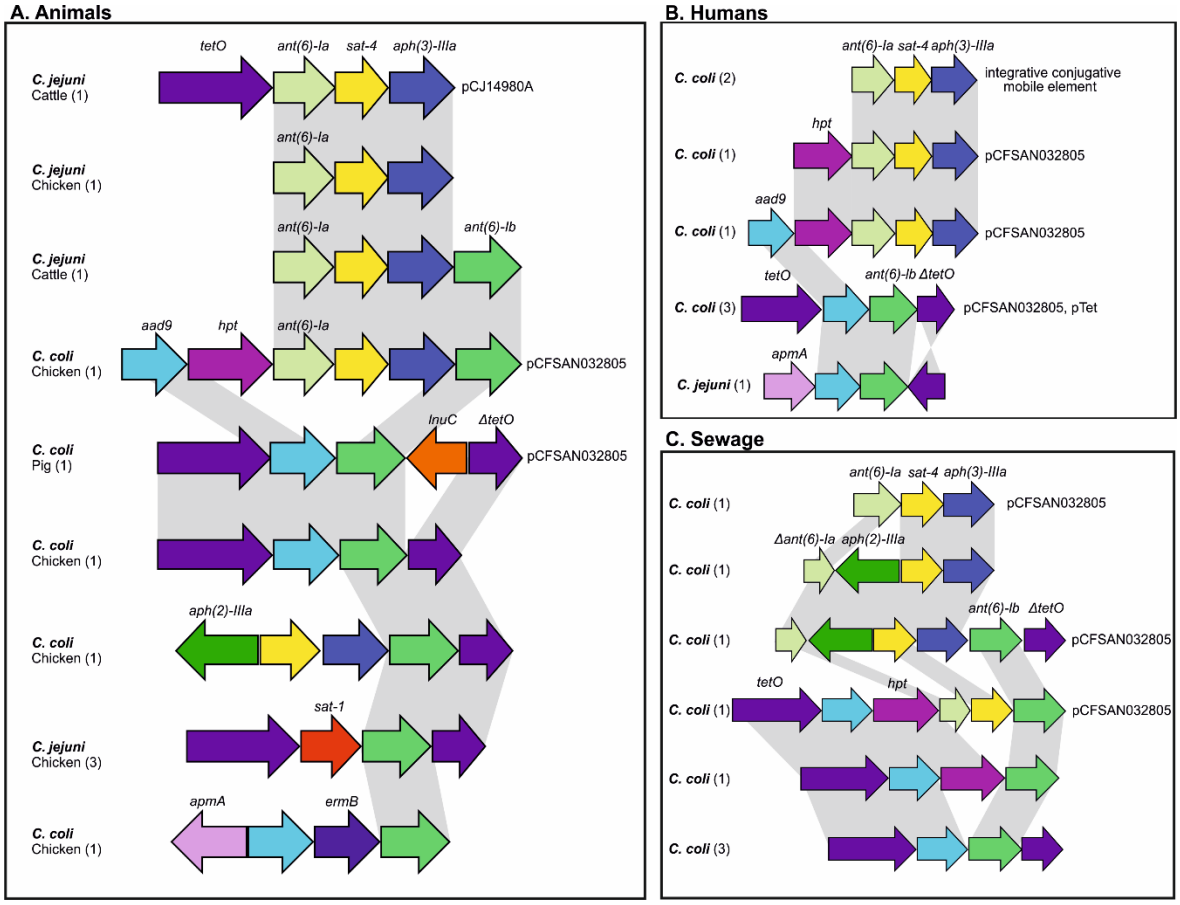


Figure 4.5. Comparative genetic organization of AMR GIs in *Campylobacter*. The presence of each AMR gene, highlighted in different colours, is shown for representative *C. jejuni* and *C. coli* isolate genomes sampled from animals (A), humans (B) and sewage (C). The number of isolate genomes containing each genomic island arrangement is indicated in the parenthesis. Grey shading identifies sequence that shares >95% nucleotide sequence identity. The name of the plasmid or mobile genetic element, associated with each genomic island, is indicated.

Clonal descent is disrupted in AMR genes

The mean consistency index (CI) was significantly higher (MannWhitney test; $U = 3307$, $p=0.0214$) among AMR genes (0.65581 ± 0.3531) compared with 595 core genes (0.4552 ± 0.05799) (Figure 4.6). This provides evidence that the clonal mode of descent has been disrupted in AMR genes consistent with HGT. Furthermore, there was a significant decrease in the average allelic variation among AMR genes compared to core genes (MannWhitney test; $U = 1004$, $p = <0.0001$) (Figure 4.7). The average number of unique alleles per isolate was 0.03436

± 0.05218 for the 15 AMR genes, compared with 0.1169 ± 0.05248 for 595 core genes. This is consistent with HGT facilitating the movement of AMR genes into multiple genetic backgrounds.

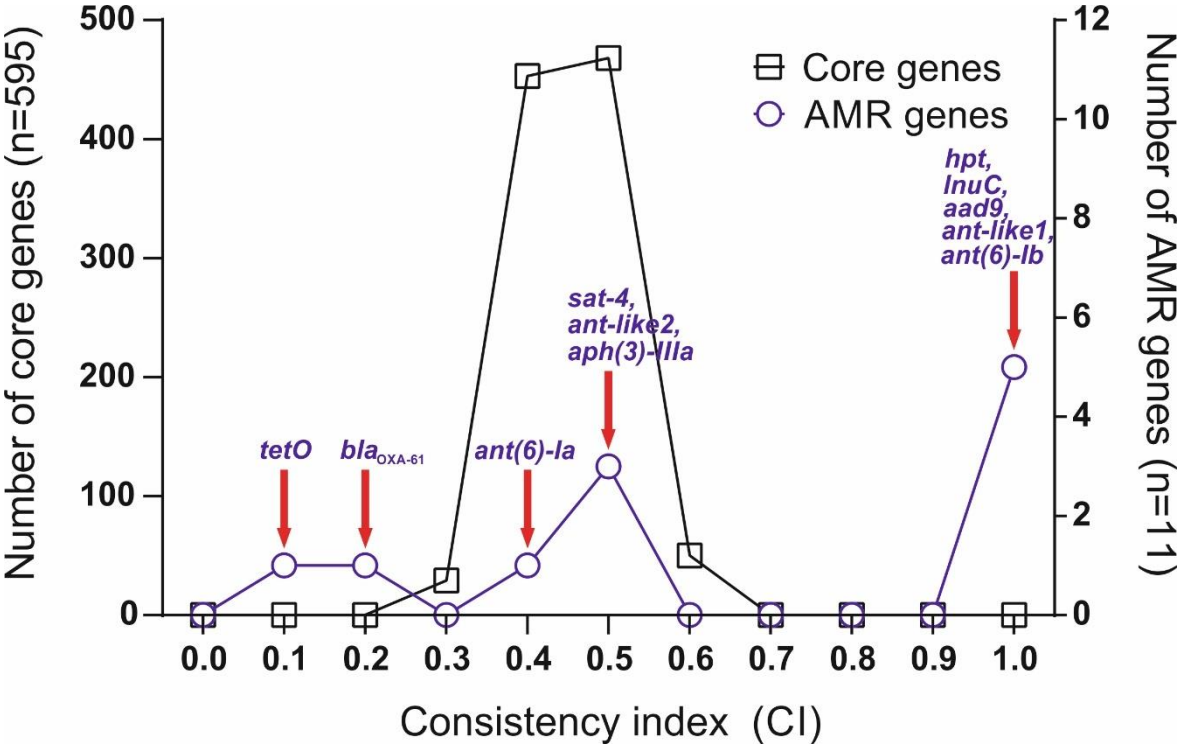


Figure 4.6. Comparison of consistency index between AMR and core genes. Consistency indices to a core phylogeny, were calculated for each gene alignment for AMR and core genes using the *phangorn* package in R. For the CI, the two distributions were significantly different (two-tailed Mann-Whitney test; $P = 0.0214$, Mann-Whitney $U = 3307$).

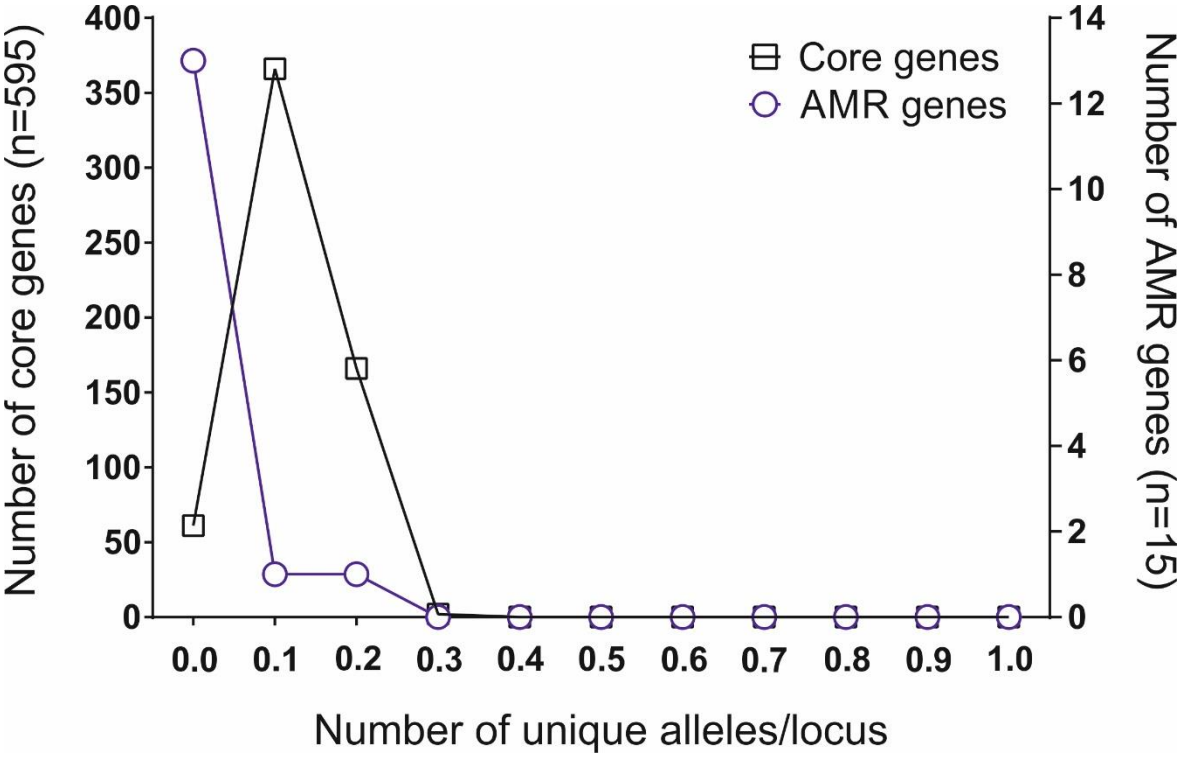


Figure 4.7. Comparison of allelic variation between AMR and core genes. The number of alleles per locus. The left *y*-axis indicates the number of core genes (black line), the right *y*-axis indicates the number of AMR genes (blue line). For the number of alleles per locus, the two distributions were significantly different (two-tailed Mann-Whitney test; $P < 0.0001$, Mann-Whitney $U = 1004$).

Among the AMR genes present in five or more isolates, the *bla*$_{\text{OXA-61}}$ and *tetO* alleles, associated with resistance to β-lactams and tetracyclines respectively, were almost ubiquitous among *C. jejuni* and in *C. coli* from different sources. Two common *bla*$_{\text{OXA-61}}$ alleles were present in both *Campylobacter* species in all different hosts and sewage with other alleles shared only between human, chicken and sewage isolates (Figure 4.8). A single *tetO* allele was present in the genomes of isolates from all different hosts and sewage except for *C. jejuni* from humans and *C. coli* cattle (Figure 4.9), possibly due to low sample numbers (Appendix, Table S4.1).

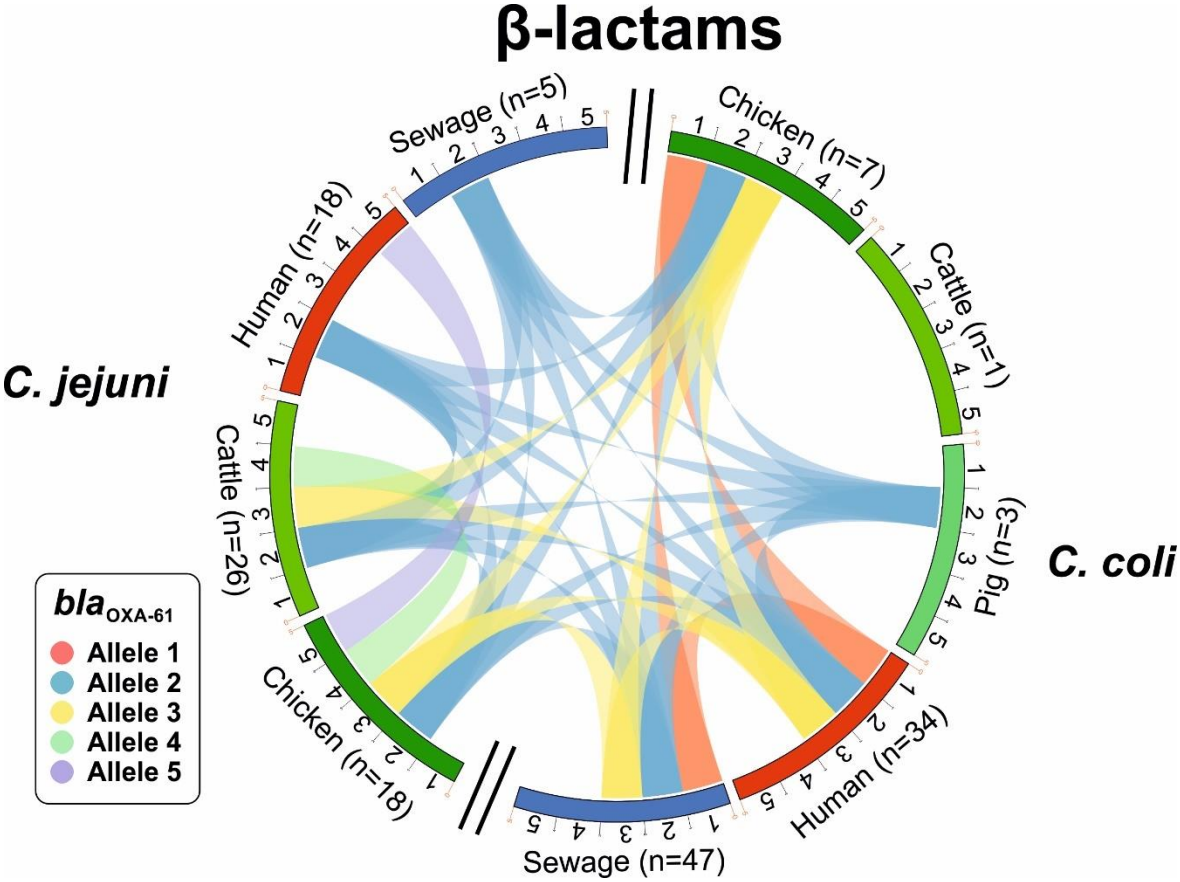


Figure 4.8. Distribution of *bla*$_{\text{OXA-61}}$ alleles among *Campylobacter* species and isolate sources. Circos plot indicate the number of *C. jejuni* and *C. coli* isolates sampled from chickens (dark green), cattle (green), pigs (light green), humans (red) and sewage (blue) that contain alleles associated with resistance to β-Lactam. Alleles present in >5 isolate genomes are numbered around the perimeter. Exact matches between allele sequences are indicated by joining lines, coloured differently for different alleles.

Another *tetO* allele was shared between *C. coli* isolates from sewage and *C. jejuni* from chickens, cattle and humans (Figure 4.9). In addition to evidence of frequent allele sharing between *Campylobacter* species from multiple sources, there were also several species-specific alleles found in isolates from multiple sources. (Figure 4.5). AMR genes associated with aminoglycoside resistance had less allelic diversity compared to *bla*$_{OXA-61}$ and *tetO* (Figure 4.3, Figure 4.4) and showed evidence of gene pool transmission between bacterial species and isolate source populations.

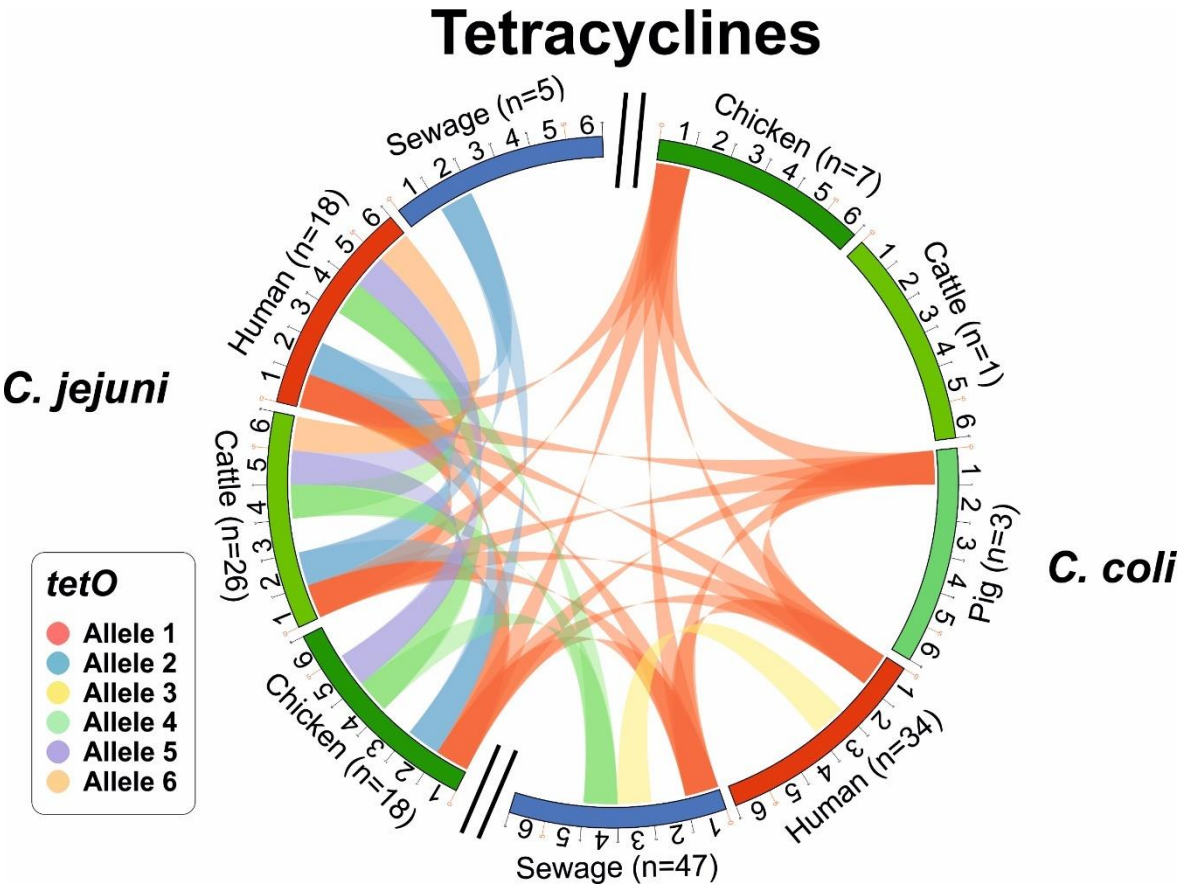


Figure 4.9. Distribution of *tetO* alleles among *Campylobacter* species and isolate sources. Circos plot indicate the number of *C. jejuni* and *C. coli* isolates sampled from chickens (dark green), cattle (green), pigs (light green), humans (red) and sewage (blue) that contain alleles associated with resistance to tetracyclines. Alleles present in >5 isolate genomes are numbered around the perimeter. Exact matches between allele sequences are indicated by joining lines, coloured differently for different alleles.

Three alleles of the *aad9, ant(6)-Ib, sat-4* genes were shared between *C. jejuni* and *C. coli* isolates. The *ant(6)-Ib* allele was found in *C. jejuni* isolates from humans, cattle, chickens and in *C. coli* isolates from humans, chickens and sewage. The *aad9* allele was found in human *C. jejuni* isolates and in *C. coli* isolates from

humans, chickens, pigs and sewage. The *sat-4* allele was found in *C. jejuni* isolates from cattle and chicken and in *C. coli* isolates from human, chicken and sewage sources (Figure 4.10). Alleles of other genes associated with aminoglycoside resistance (*ant-like* A*, aad9, ant(6)-Ib, aph(3)-IIIa, hpt* and *aph(2)-IIIa*) also showed evidence of transfer (allele sharing) between isolates sampled from different sources (Figure 4.10).

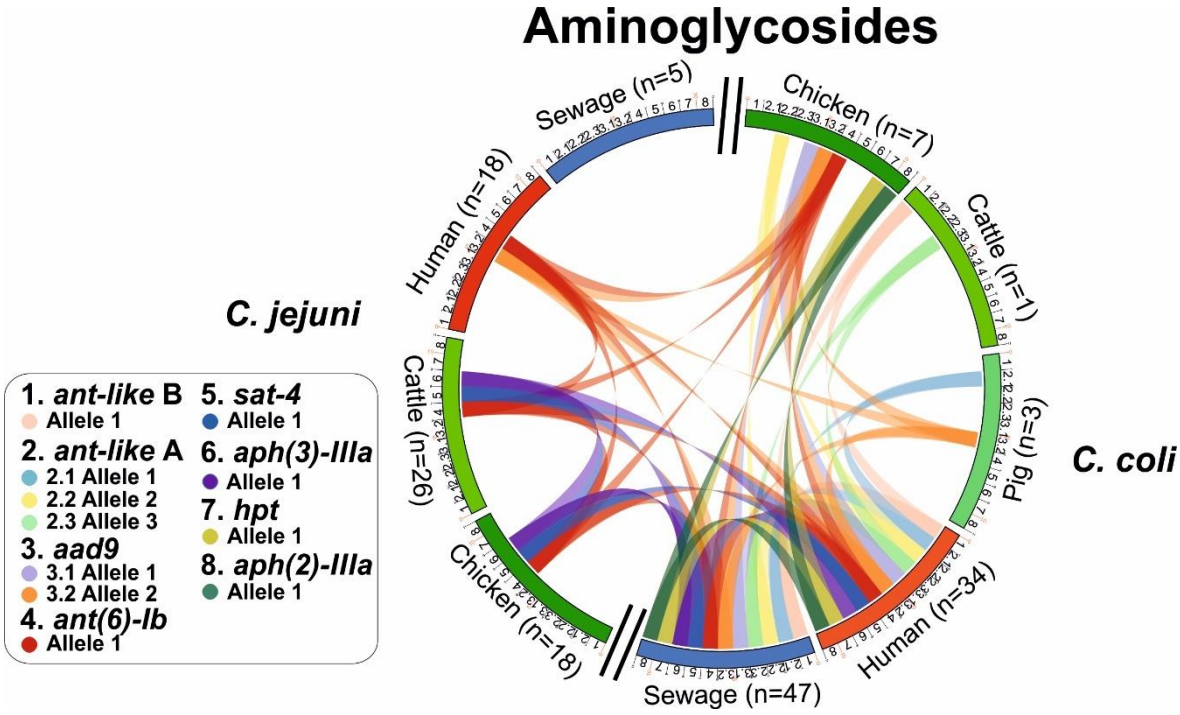


Figure 4.10. Distribution of AMR gene alleles among *Campylobacter* species and isolate sources. Circos plot indicate the number of *C. jejuni* and *C. coli* isolates sampled from chickens (dark green), cattle (green), pigs (light green), humans (red) and sewage (blue) that contain alleles associated with resistance to aminoglycosides. Alleles present in >5 isolate genomes are numbered around the perimeter. Exact matches between allele sequences are indicated by joining lines, coloured differently for different alleles.

Discussion

Forecasts of rising antimicrobial resistance in bacteria can make dramatic claims, such as an associated death toll of 10 million people by 2050 if no action is taken (Balouiri et al., 2016). However, for priority pathogens such as *Campylobacter* (WHO, 2017) it is not always clear where such action should be targeted. One reason for this is that zoonotic bacteria do not reside in a single host niche, therefore the source and sink dynamics of resistant strains may be poorly understood. Furthermore, the conduit for transmission between niches (in this case faeces) may also represent a reservoir of AMR. Here, by focussing analyses on comparison of gene pools, rather than individual resistant clones, I directly address if the alleles

that confer resistance have spread between pathogenic *Campylobacter* species and the niches in which they reside.

Human infection is often a dead-end for *Campylobacter* as disease is usually self-limiting and human-to-human transmission is uncommon. As antibiotic treatment for campylobacteriosis is generally only given in acute or persistent cases, the heavy use of related antimicrobials in human and veterinary medicine (Livermore, 2007; Schwarz et al., 2001; Teuber, 2001) has raised concerns about how selection for resistance in livestock may lead to AMR in human pathogens. Despite the ban on the use of antibiotics as growth promoters in animals in 2006, quinolones and tetracyclines are still available for treatment of livestock all over the world (WHO, 2017). Consistent with trends in a recent ECDC report (Food and Authority, 2019), resistance to ciprofloxacin and tetracycline was seen in both *Campylobacter* species in this chapter, with resistance to streptomycin and gentamycin also frequent among sequenced *C. coli* isolates (Table 4.1). This may not be surprising as Spain has the highest sale of aminoglycosides for veterinary use in the EU (European Medicines Agency, 2018). Perhaps equally worrying was the isolation of *C. coli* resistant to erythromycin which is the drug of choice for antibiotic treatment of clinical campylobacteriosis (Acheson and Allos, 2001). The extent to which this level of resistance is a legacy of past use of fluoroquinolones, tetracyclines (Cameron and Gaynor, 2014; Toth et al., 2013) and other antimicrobials is not known but it is clear that *Campylobacter* harbour numerous resistance genes, potentially exacerbated by the carriage of similar genes among other components of the microbiota (Holmes et al., 2016; van den Bogaard, 2000).

Antimicrobial resistance is widespread among *Campylobacter* isolated from livestock (Florez-Cuadrado et al., 2016; Qin et al., 2014; Wang et al., 2014; Sproston et al., 2018), but the transmission dynamics are poorly understood. Where resistance is conferred by a single (or few) nucleotide substitution(s), such as in the *gyrA* gene (fluoroquinolone resistance) (Zhao et al., 2016; Engberg et al., 2001; Payot et al., 2006), it is impossible to tell from sequence data if HGT or point mutation were responsible. For other classes of antibiotics, for example tetracyclines, there is evidence for the transfer of genes (e.g. *tetO*) between *C. jejuni* isolates, even in the absence antimicrobial selective pressure (Qin et al., 2012). In addition to *tetO*, the analyses identified 14 other accessory genes associated with

Campylobacter resistance to other known antimicrobial classes (Appendix, Table S4.4). These included aminoglycosides (10 genes), β-lactams (*bla*$_{OXA-61}$) and macrolides (*erm*B) that have been variously used as treatments targeting *Campylobacter* and other infectious agents (or even as growth promoters (Engberg et al., 2001)) in humans and animals (Lapierre et al., 2016; Qin et al., 2012; Toth et al., 2013; Lambert et al., 1985; Qin et al., 2014; Florez-Cuadrado et al., 2016; Yao et al., 2017; Chen et al., 2013; Griggs et al., 2009; Engberg et al., 2001; Florez-Cuadrado et al., 2017). Initial evidence of the importance of HGT in the transmission of these genes can be seen with inconsistent topology of individual AMR gene trees, compared to the *Campylobacter* core genome phylogeny (Appendix, Figure S4.1). Specifically, the CI varied for the 11 AMR genes, highlighting a disparity in the amount of inferred homoplasy in these genes, compared to genes in the core genome (Figure 4.6). Furthermore, the allelic variation in the AMR-associated genes was significantly lower than the mean for core genes. Convergent genotypes may have evolved multiple times in different genetic backgrounds, however the most parsimonious explanation is the spread of AMR via HGT.

Perhaps the most compelling evidence for HGT is the identification of co-localized clusters of genes that constitute GIs. Consistent with evidence of aminoglycoside resistance in *Campylobacter* (Qin et al., 2012; Gibreel et al., 2004; Lapierre et al., 2016; Lambert et al., 1985), all AMR genes detected in this chapter were found in multidrug resistance GIs, except for bla$_{OXA-61}$, *ant-like* A and *ant-like* B. There were multiple syntenic arrangements of genes with some GIs containing genes that confer resistance to more than one antimicrobial drug class (macrolides and aminoglycosides) as previously reported (Werner et al., 2003). Some of the multidrug resistance GIs are known from previous studies (*ant(6)-Ia, sat-4* and *aph(3)-IIIa*) (Derbise et al., 1997; Derbise et al., 1996), while others are reported here for the first time, such as the association between *TetO, aad9* and *ant(6)-Ib* genes. GI similarities provide evidence of transfer between *C. jejuni* and *C. coli,* and gene pool transmission among isolates from animals, humans and sewage. The transfer of GIs in *Campylobacter* can be via natural transformation (Qin et al., 2012), however several GIs were found on plasmids or integrative conjugative elements (Figure 4.5) indicating the active mobilization of gene clusters. GIs containing the *ant(6)-Ia, sat-4* and *aph(3)-IIIa* cluster, and the *tetO* gene, have previously been

described in staphylococci (Lambert et al., 1985; Derbise et al., 1996; Derbise et al., 1997). Furthermore, the conjugative transposon found in *C. coli* was highly similar (~99.4% nucleotide identity over at least 60% of the sequence) to related sequence in other Gram-positive bacteria. This is consistent with the circulation of AMR genes not only among *Campylobacter* species in different habitats but also HGT from other bacteria (Trieu-Cuot et al., 1985; Zilhao et al., 1988).

An important finding in this chapter was that *C. coli* carry more combinations of AMR genes simultaneously than *C. jejuni* (Table 4.2). A simple explanation could be that *C. coli* ST-828 complex isolates are more recombinogenic. There is evidence of the accumulation of *C. jejuni* DNA throughout the genome of this lineage (Sheppard et al., 2008; Sheppard et al., 2013) which could have led to the acquisition of multiple AMR genes. It is also possible that the dominance of this *C. coli* lineage (ST-828 complex), that is much less diverse than *C. jejuni* as a whole, reflects a genetic bottleneck that favoured an ancestral AMR strain in, for example, the pig gut where *C. coli* (Thakur et al., 2006) and antimicrobial exposure (Aarestrup et al., 2000) are common. Whatever the reasons for differences in multidrug resistance between *C. jejuni* and *C. coli,* there is clear evidence for HGT and the transmission of AMR genes among bacterial species and host niche gene pools.

Contrasting evidence of HGT with quantitative information about the transmission of resistant bacteria between hosts would be extremely useful for understanding the dissemination of AMR among isolates from different habitats. In *Campylobacter*, studies have attempted to estimate the number of strains excreted into the environment by different animals (Ogden et al., 2009) and attribute the source of human infection to livestock (especially poultry) reservoirs (Sheppard et al., 2009a; Thépault et al., 2017; Thépault et al., 2018). However, these large-scale probabilistic studies are utterly underpowered for investigating the almost infinite number of possible transmission events, where the survival and proliferation of a single strain in a new niche could lead to the transfer of AMR genes between hosts and environments. A theoretical solution to the spread of AMR could be to use different drug classes in animals on the assumption that distinct antimicrobial selection pressures would sustain efficacy of drugs in humans. However, even if this was feasible, evidence from this chapter (and others (Hendriksen et al., 2019)) shows that multidrug resistant bacteria can be isolated and cultured from sewage,

presenting a potential route for transmission of AMR in the environment. While the sources and implications of environmental contamination remain controversial (Munck et al., 2015; Rizzo et al., 2013), the evidence in this chapter is consistent with the horizontal transfer of AMR among *Campylobacter* isolated from livestock, humans and sewage. This suggests that judicious use of antimicrobials and monitoring of the amount of AMR *Campylobacter* entering the environment may be beneficial in combating the rise of resistance in this important zoonotic pathogen.

Materials and methods

Culture and antimicrobial susceptibility testing

As part of routine *Campylobacter* surveillance in Spain, isolates were sampled and cultured on blood agar plates (bioMérieux) and incubated for 48 h at 37 ℃ under microaerophilic conditions using Campygen atmosphere generation system packs (Oxoid, Basingstoke, UK). Subcultured colonies were harvested and suspended in sterile water to a standardized cell density (0.5 McFarland turbidity). 50 µL of this suspension was added to 11 ml of Mueller-Hinton broth (TREK Diagnostics Systems, Waltham, MA, USA) supplemented with 5.5% lysed horse blood (Oxoid). The solution was poured onto EUCAMP2 microdilution plates (TREK Diagnostics Systems) which were incubated under microaerophilic conditions for 48 h at 37 ℃ as previously described (Florez-Cuadrado et al., 2017). The interpretation of the quantitative data was performed according to the European Committee of Antimicrobial Susceptibility Testing, EUCAST (http://www.eucast.org/; last accessed: 06/2017). [N.B. This section of analysis was performed by Diego Flórez Cuadrado by the Universidad Complutense Madrid, Madrid, Spain.]

DNA extraction, genome sequencing and archiving

A total of 260 *Campylobacter* isolates (167 *C. jejuni* and 92 *C. coli*) that displayed multidrug resistance phenotypes were chosen for genome sequencing. These represented strains sampled from humans, livestock and urban effluents in Spain. Of these, 55 isolates originated in animals (44 *C. jejuni* and 11 *C. coli*) including broiler chickens (18 *C. jejuni* and 7 *C. coli*), cattle (26 *C. jejuni* and 1 *C. coli*) and pigs (3 *C. coli*) and were collected from abattoirs in Spain (2008-2011) as part of the Spanish Veterinary Antimicrobial Resistance Surveillance (VAV) Network (Appendix, Table S4.1). The isolates were chosen on the basis of resistance profiles (susceptible to resistant) to 5 different antibiotics (Table 4.1). Human samples

(n=152; 118 *C. jejuni* and 34 *C. coli*) were associated with campylobacteriosis cases in hospitals in the regions of Castilla y Leon, Extremadura and Andalucía between 2013 and 2016. *Campylobacter* isolates of urban effluent origin (n=53; 6 *C. jejuni* and 47 *C. coli*) were collected from the wastewater treatment plants in the city of Madrid (Spain) between 2011 and 2013 (Ugarte-Ruiz et al., 2015). All isolates were obtained using culture based methods (Moreno et al., 2000; Hormeño et al., 2016; Ugarte-Ruiz et al., 2015) and speciated as *C. jejuni* or *C. coli* using a conventional multiplex PCR as previously described (Ugarte-Ruiz et al., 2012). [N.B. This section of analysis was performed by Diego Flórez Cuadrado by the Universidad Complutense Madrid, Madrid, Spain.]

For genome sequencing, isolates stored at -80 ℃ in 1% protease peptone and 10% glycerol broth were cultured onto blood agar plates (bioMérieux) in microaerophilic conditions at 42 ℃ for 48 h as previously described (Florez-Cuadrado et al., 2017). Genomic DNA was extracted using the QIAamp DNA Mini Kit (QIAGEN, Crawley, UK), according to manufacturer's instructions. Nucleic acid content was quantified on a Nanodrop spectrophotometer prior to normalization and sequencing. Libraries were prepared with Nextera XT kits (v2) and high-throughput sequencing was performed using an Illumina MiSeq sequencer (Illumina, San Diego, CA; v3 technology, 300bp paired-end). Short reads were assembled *de novo* using SPAdes (version 3.8.0) (Bankevich et al., 2012). All genomes used in this chapter were archived on the BIGSdb web-based database platform (Jolley and Maiden, 2010) and given a unique identification number (BIGSid) (Appendix, Table S4.1).

Phylogenetic analysis

A pangenome was created for all isolate genomes in the collection as the sum of core genes, shared by all isolates, and accessory genes, present in at least one isolate. Genomes with a total assembly length >1.9Mbp, >500 contigs, or an N_{95}<800bp were considered poor quality and were excluded from the phylogenetic analyses. Whole genome multiple sequence alignments were obtained using MAFFT (Katoh, 2002) following a gene-by-gene approach as previously described (Méric et al., 2014). Phylogenetic trees, based on gene-by-gene alignments of core genes (Méric et al., 2014) or single gene sequences, were reconstructed using the Neighbor joining clustering method (Saitou and Nei, 1987). [N.B. This section of

data analysis was performed in collaboration with Diego Flórez Cuadrado by the Universidad Complutense Madrid, Madrid, Spain.]

Screening for antimicrobial resistance genes

AMR genes were identified in all *Campylobacter* genomes by comparison with the CARD (Jia et al., 2017) (last assessed: 03/06/2017), the ResFinder (Zankari et al., 2012) and the NCBI databases using the BLAST algorithm (Sheppard et al., 2012; Maiden et al., 2013). A locus match was defined when genes had >70% nucleotide identity over >50% of the sequence length, and a matrix was generated that contained presence/absence information for each card gene and the allelic variation at that locus for every genome. Following the identification of isolate genomes harbouring one or more AMR genes, contigs were screened for upstream and downstream open reading frames (ORFs) to characterize the location of AMR relative to adjacent genes, using SnapGene® software (GSL Biotech; available at snapgene.com). A second confirmatory analysis was performed, in which contigs were compared to NCBI database to identify whether they are associated with known plasmid or mobile elements. Sequence matches with >95% nucleotide identity over >50% of the sequence length were considered positive hits. A bivariate analysis was performed, in Stata version 14.0 (StataCorp, College Station, TX), to determine the relationship between phenotypes and genotypes for the presence of resistance using the Fisher's exact test. Associations were considered significant when $p<0.05$. [N.B. This section of data analysis was performed in collaboration with Diego Flórez Cuadrado by the Universidad Complutense Madrid, Madrid, Spain.]

HGT among infection-associated genes

Population genetic analyses were undertaken to compare molecular variation among AMR genes to investigate patterns of HGT between species and isolates sampled form different niches. Genes where AMR is mediated by single nucleotide polymorphisms (SNPs), for example *gyr*A in fluroquinolone resistance (Sproston et al., 2018), were excluded from this analysis because of the inability to distinguish *de novo* mutation from homologous recombination of similar sequence. The allelic variation was calculated at loci associated with AMR genes (n=15) and compared to variation at core loci (n = 595 genes). For both groups, the number of alleles at each locus (determined using MLST approach) (Sheppard et al., 2012) and CI were calculated. The consistency of a phylogenetic tree to patterns of variation in

sequence alignments was determined for each gene of interest, and constituted an inference of the minimum amount of homoplasy in these genes, as implied by the tree (Kluge and Farris, 1969). The CI function from the R Phangorn package (Schliep, 2011) was used to calculate consistency indices for every single-gene alignment of the 15 AMR genes to a phylogeny constructed from a concatenated gene-by-gene alignment of 595 core genes shared by all 259 isolates. The average CI of these shared genes was compared to that of the AMR genes.

References

Aarestrup, F.M., Kruse, H., Tast, E., Hammerum, A.M. and Jensen, L.B., 2000. Associations Between the Use of Antimicrobial Agents for Growth Promotion and the Occurrence of Resistance among Enterococcus faecium from Broilers and Pigs in Denmark, Finland, and Norway. *Microbial Drug Resistance*, 6(1), pp.63–70.

Achard, A., Villers, C., Pichereau, V. and Leclercq, R., 2005. New lnu(C) Gene Conferring Resistance to Lincomycin by Nucleotidylation in Streptococcus agalactiae UCN36. *Antimicrobial Agents and Chemotherapy*, 49(7), pp.2716–2719.

Acheson, D. and Allos, B.M., 2001. Campylobacter jejuni Infections: Update on Emerging Issues and Trends. *Clinical Infectious Diseases*, 32(8), pp.1201–1206.

Alfredson, D.A. and Korolik, V., 2005. Isolation and Expression of a Novel Molecular Class D -Lactamase, OXA-61, from Campylobacter jejuni. *Antimicrobial Agents and Chemotherapy*, 49(6), pp.2515–2518.

Atterby, C., Mourkas, E., Méric, G., Pascoe, B., Wang, H., Waldenström, J., Sheppard, S.K., Olsen, B., Järhult, J.D. and Ellström, P., 2018. The Potential of Isolation Source to Predict Colonization in Avian Hosts: A Case Study in Campylobacter jejuni Strains From Three Bird Species. *Frontiers in Microbiology*, 9(MAR).

Balouiri, M., Sadiki, M. and Ibnsouda, S.K., 2016. Methods for in vitro evaluating antimicrobial activity: A review. *Journal of Pharmaceutical Analysis*, 6(2), pp.71–79.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M.A. and Pevzner, P.A.,

2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*, 19(5), pp.455–477.

Batchelor, R.A., 2004. Nucleotide sequences and comparison of two large conjugative plasmids from different Campylobacter species. *Microbiology*, 150(10), pp.3507–3517.

Boerlin, P. and Reid-Smith, R.J., 2008. Antimicrobial resistance: its emergence and transmission. *Animal Health Research Reviews*, 9(02), pp.115–126.

van den Bogaard, A., 2000. Epidemiology of resistance to antibiotics Links between animals and humans. *International Journal of Antimicrobial Agents*, 14(4), pp.327–335.

Cameron, A. and Gaynor, E.C., 2014. Hygromycin B and Apramycin Antibiotic Resistance Cassettes for Use in Campylobacter jejuni. *PLoS ONE*, 9(4), p.e95084.

Castanon, J.I.R., 2007. History of the Use of Antibiotic as Growth Promoters in European Poultry Feeds. *Poultry Science*, 86(11), pp.2466–2471.

Chen, Y., Mukherjee, S., Hoffmann, M., Kotewicz, M.L., Young, S., Abbott, J., Luo, Y., Davidson, M.K., Allard, M., McDermott, P. and Zhao, S., 2013. Whole-Genome Sequencing of Gentamicin-Resistant Campylobacter coli Isolated from U.S. Retail Meats Reveals Novel Plasmid-Mediated Aminoglycoside Resistance Genes. *Antimicrobial Agents and Chemotherapy*, 57(11), pp.5398–5405.

Cody, A.J., McCarthy, N.D., Bray, J.E., Wimalarathna, H.M.L., Colles, F.M., Jansen van Rensburg, M.J., Dingle, K.E., Waldenström, J. and Maiden, M.C.J., 2015. Wild bird-associated C ampylobacter jejuni isolates are a consistent source of human disease, in Oxfordshire, United Kingdom. *Environmental Microbiology Reports*, 7(5), pp.782–788.

Colles, F.M., Jones, K., Harding, R.M. and Maiden, M.C.J., 2003. Genetic Diversity of Campylobacter jejuni Isolates from Farm Animals and the Farm Environment. *Applied and Environmental Microbiology*, 69(12), pp.7409–7413.

Dearlove, B.L., Cody, A.J., Pascoe, B., Méric, G., Wilson, D.J. and Sheppard, S.K., 2016. Rapid host switching in generalist Campylobacter strains erodes the signal for tracing human infections. *The ISME Journal*, 10(3), pp.721–729.

Derbise, A., de Cespedes, G. and Solh, N. El, 1997. Nucleotide sequence of

theStaphylococcus aureus transposon, Tn5405, carrying aminoglycosides resistance genes. *Journal of Basic Microbiology*, 37(5), pp.379–384.

Derbise, A., Dyke, K.G.H. and El Solh, N., 1996. Characterization of aStaphylococcus aureusTransposon, Tn5405,Located within Tn5404and Carrying the Aminoglycoside Resistance Genes,aphA-3andaadE. *Plasmid*, 35(3), pp.174–188.

Dingle, K.E., Colles, F.M., Wareing, D.R., Ure, R., Fox, A.J., Bolton, F.E., Bootsma, H.J., Willems, R.J., Urwin, R. and Maiden, M.C., 2001. Multilocus sequence typing system for Campylobacter jejuni. *Journal of clinical microbiology*, 39(1), pp.14–23.

Engberg, J., Aarestrup, F.M., Taylor, D.E., Gerner-Smidt, P. and Nachamkin, I., 2001. Quinolone and macrolide resistance in Campylobacter jejuni and C. coli: resistance mechanisms and trends in human isolates. *Emerging infectious diseases*, 7(1), pp.24–34.

European Centre for Disease Prevention and Control [ECDC] & European Food Safety Authority [EFSA], 2015. EU Summary Report on antimicrobial resistance in zoonotic and indicator bacteria from humans, animals and food in 2013. *EFSA Journal*, 13(2).

European Medicines Agency, 2018. Reflection paper on use of aminoglycosides in animals in the European Union: development of resistance and impact on human and animal health. 44, pp.1–42.

Florez-Cuadrado, D., Moreno, M.A., Ugarte-Ruíz, M. and Domínguez, L., 2018. Antimicrobial Resistance in the Food Chain in the European Union. *Advances in Food and Nutrition Research*. pp.115–136.

Florez-Cuadrado, D., Ugarte-Ruiz, M., Meric, G., Quesada, A., Porrero, M.C., Pascoe, B., Sáez-Llorente, J.L., Orozco, G.L., Domínguez, L. and Sheppard, S.K., 2017. Genome Comparison of Erythromycin Resistant Campylobacter from Turkeys Identifies Hosts and Pathways for Horizontal Spread of erm(B) Genes. *Frontiers in Microbiology*, 8(NOV), pp.1–8.

Florez-Cuadrado, D., Ugarte-Ruiz, M., Quesada, A., Palomo, G., Domínguez, L. and Porrero, M.C., 2016. Description of an erm (B)-carrying Campylobacter coli isolate in Europe. *Journal of Antimicrobial Chemotherapy*, 71(3), pp.841–843.

Food, E. and Authority, S., 2019. The European Union summary report on antimicrobial resistance in zoonotic and indicator bacteria from humans,

animals and food in 2017. *EFSA Journal*, 17(2).

Fravalo, P., Laisney, M.-J., Gillard, M.-O., Salvat, G. and Chemaly, M., 2009. Campylobacter transfer from naturally contaminated chicken thighs to cutting boards is inversely related to initial load. *Journal of food protection*, 72(9), pp.1836–40.

Gibreel, A., 2006. Macrolide resistance in Campylobacter jejuni and Campylobacter coli. *Journal of Antimicrobial Chemotherapy*, 58(2), pp.243–255.

Gibreel, A., Sköld, O. and Taylor, D.E., 2004. Characterization of Plasmid-Mediated aphA - 3 Kanamycin Resistance in Campylobacter jejuni. *Microbial Drug Resistance*, 10(2), pp.98–105.

Griekspoor, P., Colles, F.M., McCarthy, N.D., Hansbro, P.M., Ashhurst-Smith, C., Olsen, B., Hasselquist, D., Maiden, M.C.J. and Waldenström, J., 2013. Marked host specificity and lack of phylogeographic population structure of Campylobacter jejuni in wild birds. *Molecular Ecology*, 22(5), pp.1463–1472.

Griggs, D.J., Peake, L., Johnson, M.M., Ghori, S., Mott, A. and Piddock, L.J. V., 2009. -Lactamase-Mediated -Lactam Resistance in Campylobacter Species: Prevalence of Cj0299 (blaOXA-61) and Evidence for a Novel -Lactamase in C. jejuni. *Antimicrobial Agents and Chemotherapy*, 53(8), pp.3357–3364.

Guyard-Nicodème, M., Tresse, O., Houard, E., Jugiau, F., Courtillon, C., El Manaa, K., Laisney, M.-J. and Chemaly, M., 2013. Characterization of Campylobacter spp. transferred from naturally contaminated chicken legs to cooked chicken slices via a cutting board. *International Journal of Food Microbiology*, 164(1), pp.7–14.

Hendriksen, R.S., Munk, P., Njage, P., van Bunnik, B., McNally, L., Lukjancenko, O., Röder, T., Nieuwenhuijse, D., Pedersen, S.K., Kjeldgaard, J., Kaas, R.S., Clausen, P.T.L.C., Vogt, J.K., Leekitcharoenphon, P., van de Schans, M.G.M., Zuidema, T., de Roda Husman, A.M., Rasmussen, S., Petersen, B., Amid, C., Cochrane, G., Sicheritz-Ponten, T., Schmitt, H., Alvarez, J.R.M., Aidara-Kane, A., Pamp, S.J., Lund, O., Hald, T., Woolhouse, M., Koopmans, M.P., Vigre, H., Petersen, T.N. and Aarestrup, F.M., 2019. Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nature Communications*, 10(1), p.1124.

Hermans, D., Pasmans, F., Messens, W., Martel, A., Van Immerseel, F., Rasschaert, G., Heyndrickx, M., Van Deun, K. and Haesebrouck, F., 2012.

Poultry as a Host for the Zoonotic Pathogen Campylobacter jejuni. *Vector-Borne and Zoonotic Diseases*, 12(2), pp.89–98.

Holmes, A.H., Moore, L.S.P., Sundsfjord, A., Steinbakk, M., Regmi, S., Karkey, A., Guerin, P.J. and Piddock, L.J. V, 2016. Understanding the mechanisms and drivers of antimicrobial resistance. *The Lancet*, 387(10014), pp.176–187.

Hormeño, L., Palomo, G., Ugarte-Ruiz, M., Porrero, M.C., Borge, C., Vadillo, S., Píriz, S., Domínguez, L., Campos, M.J. and Quesada, A., 2016. Identification of the main quinolone resistance determinant in Campylobacter jejuni and Campylobacter coli by MAMA-DEG PCR. *Diagnostic Microbiology and Infectious Disease*, 84(3), pp.236–239.

Hormeño, L., Ugarte-Ruiz, M., Palomo, G., Borge, C., Florez-Cuadrado, D., Vadillo, S., Píriz, S., Domínguez, L., Campos, M.J. and Quesada, A., 2018. ant(6)-I Genes Encoding Aminoglycoside O-Nucleotidyltransferases Are Widely Spread Among Streptomycin Resistant Strains of Campylobacter jejuni and Campylobacter coli. *Frontiers in Microbiology*, 9(OCT), pp.1–8.

Huttner, A., Harbarth, S., Carlet, J., Cosgrove, S., Goossens, H., Holmes, A., Jarlier, V., Voss, A. and Pittet, D., 2013. Antimicrobial resistance: a global view from the 2013 World Healthcare-Associated Infections Forum. *Antimicrobial Resistance and Infection Control*, 2(1), p.31.

Iovine, N.M., 2013. Resistance mechanisms in Campylobacter jejuni. *Virulence*, 4(3), pp.230–240.

Jia, B., Raphenya, A.R., Alcock, B., Waglechner, N., Guo, P., Tsang, K.K., Lago, B.A., Dave, B.M., Pereira, S., Sharma, A.N., Doshi, S., Courtot, M., Lo, R., Williams, L.E., Frye, J.G., Elsayegh, T., Sardar, D., Westman, E.L., Pawlowski, A.C., Johnson, T.A., Brinkman, F.S.L., Wright, G.D. and McArthur, A.G., 2017. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Research*, 45(D1), pp.D566–D573.

Jolley, K.A. and Maiden, M.C.J., 2010. BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC bioinformatics*, 11(1), p.595.

Katoh, K., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30(14), pp.3059–3066.

Kluge, A.G. and Farris, J.S., 1969. Quantitative Phyletics and the Evolution of Anurans. *Systematic Biology*, 18(1), pp.1–32.

Lambert, T., Gerbaud, G., Trieu-Cuot, P. and Courvalin, P., 1985. Structural relationship between the genes encoding 3′-aminoglycoside phosphotransferases in Campylobacter and in gram-positive cocci. *Annales de l'Institut Pasteur / Microbiologie*, 136(2), pp.135–150.

Lapierre, L., Arias, M.L. and Fernández, H., 2016. Antimicrobial Resistance in Campylobacter spp. *Campylobacter spp. and Related Organisms in Poultry*. Cham: Springer International Publishing, pp.165–183.

Li, J., Jiang, N., Ke, Y., Feßler, A.T., Wang, Y., Schwarz, S. and Wu, C., 2017. Characterization of pig-associated methicillin-resistant Staphylococcus aureus. *Veterinary Microbiology*, 201, pp.183–187.

Livermore, D.M., 2007. Introduction: the challenge of multiresistance. *International Journal of Antimicrobial Agents*, 29(SUPPL. 3), pp.S1–S7.

Luangtongkum, T., Jeon, B., Han, J., Plummer, P., Logue, C.M. and Zhang, Q., 2009. Antibiotic resistance in Campylobacter : emergence, transmission and persistence. *Future Microbiology*, 4(2), pp.189–200.

Luo, N., Sahin, O., Lin, J., Michel, L.O. and Zhang, Q., 2003. In Vivo Selection of Campylobacter Isolates with High Levels of Fluoroquinolone Resistance Associated with gyrA Mutations and the Function of the CmeABC Efflux Pump. *Antimicrobial Agents and Chemotherapy*, 47(1), pp.390–394.

Maiden, M.C.J., van Rensburg, M.J.J., Bray, J.E., Earle, S.G., Ford, S.A., Jolley, K.A. and McCarthy, N.D., 2013. MLST revisited: the gene-by-gene approach to bacterial genomics. *Nature Reviews Microbiology*, 11(10), pp.728–736.

McDermott, P.F., Tyson, G.H., Kabera, C., Chen, Y., Li, C., Folster, J.P., Ayers, S.L., Lam, C., Tate, H.P. and Zhao, S., 2016. Whole-Genome Sequencing for Detecting Antimicrobial Resistance in Nontyphoidal Salmonella. *Antimicrobial Agents and Chemotherapy*, 60(9), pp.5515–5520.

Méric, G., Yahara, K., Mageiros, L., Pascoe, B., Maiden, M.C.J., Jolley, K.A. and Sheppard, S.K., 2014. A Reference Pan-Genome Approach to Comparative Bacterial Genomics: Identification of Novel Epidemiological Markers in Pathogenic Campylobacter. *PLoS ONE*, 9(3), p.e92798.

Miller, W.G., 2006. Identification of host-associated alleles by multilocus sequence typing of Campylobacter coli strains from food animals. *Microbiology*, 152(1), pp.245–255.

Moreno, M.A., Domínguez, L., Teshager, T., Herrero, I. a and Porrero, M.C., 2000. Antibiotic resistance monitoring: the Spanish programme. The VAV Network.

Red de Vigilancia de Resistencias Antibióticas en Bacterias de Origen Veterinario. *International journal of antimicrobial agents*, 14(4), pp.285–90.

Munck, C., Albertsen, M., Telke, A., Ellabaan, M., Nielsen, P.H. and Sommer, M.O.A., 2015. Limited dissemination of the wastewater treatment plant core resistome. *Nature Communications*, 6(1), p.8452.

Nachamkin, I., Engberg, J. and Aarestrup, F.M., 2000. Diagnosis and Antimicrobial Susceptibility of Campylobacter species. In: I. Nachamkin and M.J. Blaser, eds. [Online]. Washington DC, USA: ASM Press, pp.45–66.

Ogden, I.D., Dallas, J.F., MacRae, M., Rotariu, O., Reay, K.W., Leitch, M., Thomson, A.P., Sheppard, S.K., Maiden, M., Forbes, K.J. and Strachan, N.J.C., 2009. Campylobacter Excreted into the Environment by Animal Sources: Prevalence, Concentration Shed, and Host Association. *Foodborne Pathogens and Disease*, 6(10), pp.1161–1170.

Olkkola, S., Culebro, A., Juntunen, P., Hänninen, M.-L. and Rossi, M., 2016. Functional genomics in Campylobacter coli identified a novel streptomycin resistance gene located in a hypervariable genomic region. *Microbiology*, 162(7), pp.1157–1166.

Pascoe, B., Méric, G., Yahara, K., Wimalarathna, H., Murray, S., Hitchings, M.D., Sproston, E.L., Carrillo, C.D., Taboada, E.N., Cooper, K.K., Huynh, S., Cody, A.J., Jolley, K.A., Maiden, M.C.J., McCarthy, N.D., Didelot, X., Parker, C.T. and Sheppard, S.K., 2017. Local genes for local bacteria: Evidence of allopatry in the genomes of transatlantic Campylobacter populations. *Molecular Ecology*, 26(17), pp.4497–4508.

Payot, S., Bolla, J.-M., Corcoran, D., Fanning, S., Mégraud, F. and Zhang, Q., 2006. Mechanisms of fluoroquinolone and macrolide resistance in Campylobacter spp. *Microbes and Infection*, 8(7), pp.1967–1971.

Qin, S., Wang, Y., Zhang, Q., Chen, X., Shen, Z., Deng, F., Wu, C. and Shen, J., 2012. Identification of a Novel Genomic Island Conferring Resistance to Multiple Aminoglycoside Antibiotics in Campylobacter coli. *Antimicrobial Agents and Chemotherapy*, 56(10), pp.5332–5339.

Qin, S., Wang, Y., Zhang, Q., Zhang, M., Deng, F., Shen, Z., Wu, C., Wang, S., Zhang, J. and Shen, J., 2014. Report of ribosomal RNA methylase gene erm(B) in multidrug-resistant Campylobacter coli. *Journal of Antimicrobial Chemotherapy*, 69(4), pp.964–968.

Rizzo, L., Manaia, C., Merlin, C., Schwartz, T., Dagot, C., Ploy, M.C., Michael, I.

and Fatta-Kassinos, D., 2013. Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: A review. *Science of The Total Environment*, 447, pp.345–360.

Saitou, N. and Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4(4), pp.406–425.

Schliep, K.P., 2011. phangorn: phylogenetic analysis in R. *Bioinformatics*, 27(4), pp.592–593.

Schwarz, S., Kehrenberg, C. and Walsh, T.R., 2001. Use of antimicrobial agents in veterinary medicine and food animal production. *International Journal of Antimicrobial Agents*, 17(6), pp.431–437.

Sheppard, S.K., Cheng, L., Méric, G., de Haan, C.P.A., Llarena, A.-K., Marttinen, P., Vidal, A., Ridley, A., Clifton-Hadley, F., Connor, T.R., Strachan, N.J.C., Forbes, K., Colles, F.M., Jolley, K.A., Bentley, S.D., Maiden, M.C.J., Hänninen, M.-L., Parkhill, J., Hanage, W.P. and Corander, J., 2014. Cryptic ecology among host generalist Campylobacter jejuni in domestic animals. *Molecular Ecology*, 23(10), pp.2442–2451.

Sheppard, S.K., Colles, F., Richardson, J., Cody, A.J., Elson, R., Lawson, A., Brick, G., Meldrum, R., Little, C.L., Owen, R.J., Maiden, M.C.J. and McCarthy, N.D., 2010a. Host Association of Campylobacter Genotypes Transcends Geographic Variation. *Applied and Environmental Microbiology*, 76(15), pp.5269–5277.

Sheppard, S.K., Colles, F.M., McCarthy, N.D., Strachan, N.J.C., Ogden, I.D., Forbes, K.J., Dallas, J.F. and Maiden, M.C.J., 2011. Niche segregation and genetic structure of Campylobacter jejuni populations from wild and agricultural host species. *Molecular ecology*, 20(16), pp.3484–90.

Sheppard, S.K., Dallas, J.F., Strachan, N.J.C., MacRae, M., McCarthy, N.D., Wilson, D.J., Gormley, F.J., Falush, D., Ogden, I.D., Maiden, M.C.J. and Forbes, K.J., 2009a. Campylobacter Genotyping to Determine the Source of Human Infection. *Clinical Infectious Diseases*, 48(8), pp.1072–1078.

Sheppard, S.K., Dallas, J.F., Wilson, D.J., Strachan, N.J.C., McCarthy, N.D., Jolley, K.A., Colles, F.M., Rotariu, O., Ogden, I.D., Forbes, K.J. and Maiden, M.C.J., 2010b. Evolution of an Agriculture-Associated Disease Causing Campylobacter coli Clade: Evidence from National Surveillance Data in Scotland. *PLoS ONE*, 5(12), p.e15708.

Sheppard, S.K., Didelot, X., Jolley, K.A., Darling, A.E., Pascoe, B., Meric, G., Kelly, D.J., Cody, A., Colles, F.M., Strachan, N.J.C., Ogden, I.D., Forbes, K., French, N.P., Carter, P., Miller, W.G., McCarthy, N.D., Owen, R., Litrup, E., Egholm, M., Affourtit, J.P., Bentley, S.D., Parkhill, J., Maiden, M.C.J. and Falush, D., 2013. Progressive genome-wide introgression in agricultural Campylobacter coli. *Molecular Ecology*, 22(4), pp.1051–1064.

Sheppard, S.K., Jolley, K.A. and Maiden, M.C.J., 2012. A Gene-By-Gene Approach to Bacterial Population Genomics: Whole Genome MLST of Campylobacter. *Genes*, 3(2), pp.261–277.

Sheppard, S.K. and Maiden, M.C.J., 2015. The Evolution of Campylobacter jejuni and Campylobacter coli. *Cold Spring Harbor Perspectives in Biology*, 7(8), p.a018119.

Sheppard, S.K., McCarthy, N.D., Falush, D. and Maiden, M.C.J., 2008. Convergence of Campylobacter Species: Implications for Bacterial Evolution. *Science*, 320(5873), pp.237–239.

Sheppard, S.K.K., Dallas, J.F., MacRae, M., McCarthy, N.D., Sproston, E.L., Gormley, F.J., Strachan, N.J.C., Ogden, I.D., Maiden, M.C.J. and Forbes, K.J., 2009b. Campylobacter genotypes from food animals, environmental sources and clinical disease in Scotland 2005/6. *International Journal of Food Microbiology*, 134(1–2), pp.96–103.

Sougakoff, W., Papadopoulou, B., Nordmann, P. and Courvalin, P., 1987. Nucleotide sequence and distribution of gene tetO encoding tetracycline resistance in Campylobacter coli. *FEMS Microbiology Letters*, 44(1), pp.153–159.

Sproston, E.L., Ogden, I.D., MacRae, M., Dallas, J.F., Sheppard, S.K., Cody, A.J., Colles, F.M., Wilson, M.J., Forbes, K.J. and Strachan, N.J.C., 2011. Temporal Variation and Host Association in the Campylobacter Population in a Longitudinal Ruminant Farm Study. *Applied and Environmental Microbiology*, 77(18), pp.6579–6586.

Sproston, E.L., Wimalarathna, H.M.L. and Sheppard, S.K., 2018. Trends in fluoroquinolone resistance in Campylobacter. *Microbial Genomics*, 4(8), pp.1–8.

Taylor, D.E., 1986. Plasmid-mediated tetracycline resistance in Campylobacter jejuni: expression in Escherichia coli and identification of homology with streptococcal class M determinant. *Journal of Bacteriology*, 165(3), pp.1037–

1039.

Taylor, D.E., Garner, R.S. and Allan, B.J., 1983. Characterization of tetracycline resistance plasmids from Campylobacter jejuni and Campylobacter coli. *Antimicrobial Agents and Chemotherapy*, 24(6), pp.930–935.

Teuber, M., 2001. Veterinary use and antibiotic resistance. *Current Opinion in Microbiology*, 4(5), pp.493–499.

Thakur, S., Morrow, W.E.M., Funk, J.A., Bahnson, P.B. and Gebreyes, W.A., 2006. Molecular Epidemiologic Investigation of Campylobacter coli in Swine Production Systems, Using Multilocus Sequence Typing. *Applied and Environmental Microbiology*, 72(8), pp.5666–5669.

Thépault, A., Méric, G., Rivoal, K., Pascoe, B., Mageiros, L., Touzain, F., Rose, V., Béven, V., Chemaly, M. and Sheppard, S.K., 2017. Genome-Wide Identification of Host-Segregating Epidemiological Markers for Source Attribution in Campylobacter jejuni. *Applied and Environmental Microbiology*, 83(7), pp.1–13.

Thépault, A., Rose, V., Quesne, S., Poezevara, T., Béven, V., Hirchaud, E., Touzain, F., Lucas, P., Méric, G., Mageiros, L., Sheppard, S.K., Chemaly, M. and Rivoal, K., 2018. Ruminant and chicken: important sources of campylobacteriosis in France despite a variation of source attribution in 2009 and 2015. *Scientific Reports*, 8(1), p.9305.

Toth, M., Frase, H., Antunes, N.T. and Vakulenko, S.B., 2013. Novel Aminoglycoside 2″-Phosphotransferase Identified in a Gram-Negative Pathogen. *Antimicrobial Agents and Chemotherapy*, 57(1), pp.452–457.

Trieu-Cuot, P., Gerbaud, G., Lambert, T. and Courvalin, P., 1985. In vivo transfer of genetic information between gram-positive and gram-negative bacteria. *The EMBO Journal*, 4(13A), pp.3583–3587.

Tyson, G.H., McDermott, P.F., Li, C., Chen, Y., Tadesse, D.A., Mukherjee, S., Bodeis-Jones, S., Kabera, C., Gaines, S.A., Loneragan, G.H., Edrington, T.S., Torrence, M., Harhay, D.M. and Zhao, S., 2015. WGS accurately predicts antimicrobial resistance in Escherichia coli. *Journal of Antimicrobial Chemotherapy*, 70(10), pp.2763–2769.

Ugarte-Ruiz, M., Florez-Cuadrado, D., Wassenaar, T., Porrero, M. and Domínguez, L., 2015. Method Comparison for Enhanced Recovery, Isolation and Qualitative Detection of C. jejuni and C. coli from Wastewater Effluent Samples. *International Journal of Environmental Research and Public Health*,

12(3), pp.2749–2764.

Ugarte-Ruiz, M., Gómez-Barrero, S., Porrero, M.C., Álvarez, J., García, M., Comerón, M.C., Wassenaar, T.M. and Domínguez, L., 2012. Evaluation of four protocols for the detection and isolation of thermophilic Campylobacter from different matrices. *Journal of Applied Microbiology*, 113(1), pp.200–208.

Wang, Y., Zhang, M., Deng, F., Shen, Z., Wu, C., Zhang, J., Zhang, Q. and Shen, J., 2014. Emergence of Multidrug-Resistant Campylobacter Species Isolates with a Horizontally Acquired rRNA Methylase. *Antimicrobial Agents and Chemotherapy*, 58(9), pp.5405–5412.

Werner, G., Hildebrandt, B. and Witte, W., 2003. Linkage of erm (B) and aadE-sat4-aphA-3 in Multiple-Resistant Enterococcus faecium Isolates of Different Ecological Origins. *Microbial Drug Resistance*, 9(supplement 1), pp.9–16.

WHO, 2017. *WHO publishes list of bacteria for which new antibiotics are urgently needed*.

Wilson, D.J., Gabriel, E., Leatherbarrow, A.J.H., Cheesbrough, J., Gee, S., Bolton, E., Fox, A., Hart, C.A., Diggle, P.J. and Fearnhead, P., 2009. Rapid Evolution and the Importance of Recombination to the Gastroenteric Pathogen Campylobacter jejuni. *Molecular Biology and Evolution*, 26(2), pp.385–397.

Yahara, K., Didelot, X., Ansari, M.A., Sheppard, S.K. and Falush, D., 2014. Efficient Inference of Recombination Hot Regions in Bacterial Genomes. *Molecular Biology and Evolution*, 31(6), pp.1593–1605.

Yahara, K., Didelot, X., Jolley, K.A., Kobayashi, I., Maiden, M.C.J., Sheppard, S.K. and Falush, D., 2016. The Landscape of Realized Homologous Recombination in Pathogenic Bacteria. *Molecular Biology and Evolution*, 33(2), pp.456–471.

Yahara, K., Méric, G., Taylor, A.J., de Vries, S.P.W., Murray, S., Pascoe, B., Mageiros, L., Torralbo, A., Vidal, A., Ridley, A., Komukai, S., Wimalarathna, H., Cody, A.J., Colles, F.M., McCarthy, N., Harris, D., Bray, J.E., Jolley, K.A., Maiden, M.C.J., Bentley, S.D., Parkhill, J., Bayliss, C.D., Grant, A., Maskell, D., Didelot, X., Kelly, D.J. and Sheppard, S.K., 2017. Genome-wide association of functional traits linked with C ampylobacter jejuni survival from farm to fork. *Environmental Microbiology*, 19(1), pp.361–380.

Yao, H., Liu, D., Wang, Y., Zhang, Q. and Shen, Z., 2017. High Prevalence and Predominance of the aph(2″)-If Gene Conferring Aminoglycoside Resistance in Campylobacter. *Antimicrobial Agents and Chemotherapy*, 61(5).

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., Aarestrup, F.M. and Larsen, M.V., 2012. Identification of acquired antimicrobial resistance genes. *Journal of Antimicrobial Chemotherapy*, 67(11), pp.2640–2644.

Zhao, S., Tyson, G.H., Chen, Y., Li, C., Mukherjee, S., Young, S., Lam, C., Folster, J.P., Whichard, J.M. and McDermott, P.F., 2016. Whole-Genome Sequencing Analysis Accurately Predicts Antimicrobial Resistance Phenotypes in Campylobacter spp. *Applied and Environmental Microbiology*, 82(2), pp.459–466.

Zilhao, R., Papadopoulou, B. and Courvalin, P., 1988. Occurrence of the Campylobacter resistance gene tetO in Enterococcus and Streptococcus spp. *Antimicrobial Agents and Chemotherapy*, 32(12), pp.1793–1796.

Data availability

All sequence data are linked to NCBI BioProject PRJNA528879. The bacterial genomes are available in GenBank under accession codes SRX5575129 to SRX5587545. Contiguous assemblies of all genome sequences compared are available at the public data repository FigShare (doi.org/10.6084/m9.figshare.12743327).

Copyright agreement

Chapter 5

The potential of isolation source to predict colonization in avian hosts: a case study in *Campylobacter jejuni* strains from three bird species

The work presented in this chapter is published as a peer-reviewed publication at:

Atterby, C., Mourkas, E., Méric, G., Pascoe, B., Waldenström, J., Sheppard, S.K., Olsen, B., Järhult, J.D., Ellström, P. 2018. "The potential of isolation source to predict colonization in avian hosts: a case study in *Campylobacter jejuni* strains from three bird species." *Frontiers in Microbiology*. doi:10.3389/fmicb.2018.00591.

Commentary text

The previous three chapters have shown the existence of accessory genes that are unique in distinct *Campylobacter* species and lineages. The work in this chapter utilizes comparative genomics in order to identify molecular targets for monitoring bacterial load in *in vivo* animal infection experiments. Comparative genomics of a group of *C. jejuni* isolates was used to identified genes that are present in strains used in animal infection experiments. These genes were used as molecular targets, for a singleplex qPCR, to differentiate the DNA of strains used in competition animal experiments. Results in this chapter show the power of using genomics in a wide range of experiments and the accuracy of quantifying the DNA relative to CFU plate counts of live bacteria. The statement of authorship for this chapter is available in the Appendix, supplementary form SF4.

Abstract

Campylobacter jejuni is the primary cause of bacterial gastroenteritis worldwide, infecting humans mostly through consumption of contaminated poultry. *C. jejuni* is common in the gut of wild birds and shows distinct strain-specific association to particular bird species. This contrasts with farm animals, in which several genotypes co-exist. It is unclear if the barriers restricting transmission between host species of such specialist strains are related to environmental factors such as contact between host species, bacterial survival in the environment etc. or rather to strain specific adaptation to the intestinal environment of specific hosts. This work compared colonization dynamics *in vivo* between two host-specific *C. jejuni* from a song thrush (ST-1304 complex) and a mallard (ST-995), and a generalist strain from chicken (ST-21 complex) in a wild host, the mallard (*Anas platyrhynchos)*. In 18-day infection experiments, the song thrush strain showed only weak colonization and was cleared from all birds after 10 days, whereas both mallard and chicken strains remained stable. When the chicken strain was given 4 days prior to co-infection of the same birds with a mallard strain, it was rapidly outcompeted by the latter. In contrast, when the mallard strain was given 4 days prior to co-infection with the chicken strain, the mallard strain remained, and expansion of the chicken strain was delayed. This work's results suggest strain-specific differences in the ability of *C. jejuni* to colonize mallards, likely associated with host origin. This difference might explain observed host association patterns in *C. jejuni* from wild birds.

Introduction

The majority of human pathogens are zoonotic and able to infect more than one host species (Taylor et al., 2001; Woolhouse et al., 2001), including diseases of significant health concerns such as Salmonellosis, Tuberculosis, Cholera and Lyme disease. Furthermore, host-restricted pathogens are believed to have evolved from ancestors with a generalist life style and in some cases, this has been associated with a change in pathogenicity (Baumler and Fang, 2013). One example is *Salmonella enterica* serovar Typhi. In contrast to most of the related serovars in the *S. enterica* subspecies that are generalist enteric pathogens, *S. Typhi* separated 10,000 to 70,000 years ago to become a specialist pathogen of humans causing disseminated septicaemia (typhoid fever) (Selander et al., 1990; Roumagnac et al., 2006). The mechanisms behind host specificity for bacterial pathogens are multifactorial and include colonization, replication in the host and competition with

the surrounding microbiota (Zahrt, 1998; Baumler and Fang, 2013). In some bacterial species, specific genomic alterations are associated with specialist pathogen lineages, reviewed in Bäumler and Fang (Baumler and Fang, 2013). Such signatures can involve genomic decay and genomic rearrangements, the result of the accumulation of mutations or rearrangements of genes in the absence of selection pressure to maintain gene function. Lateral gene transfer between bacterial strains or species, can also result in the accumulation of pathogen specific genetic elements that, for example, allow the bacterium to use alternative transmission/infection routes (e.g. acquired binding to new cell types) or adaptation to the host.

An example of a multi-host zoonotic pathogen is *Campylobacter jejuni*, the leading cause of bacterial gastroenteritis in almost all industrialized countries (Food et al., 2014). *C. jejuni* has a broad host range and has been isolated from domestic (Whiley et al., 2013) and wild mammals (Petersen et al., 2001) and several bird species (Kapperud and Rosef, 1983; Waldenström et al., 2002; Colles et al., 2008a; Colles et al., 2008b). It is frequently detected in environmental waters and can even survive in unicellular eukaryotes such as amoebae (Axelsson-Olsson et al., 2005; Brennhovd et al., 1992). The most important transmission route to humans is consumption of contaminated or undercooked food items, especially from poultry (Dingle et al., 2002). Other sources of human *C. jejuni* infections are water, dairy products and other farm animals, but although the bacterium has several wild animal hosts, the extent of transmission to humans from such sources are less well studied. Interestingly, chickens are asymptomatically colonized with *C. jejuni*, suggesting commensal adaptations to the chicken gut (Humphrey et al., 2007).

Genetic relatedness and source attribution of *C. jejuni* has been studied using multilocus sequence typing (MLST). This sequence based typing approach allows clustering of genotypes into sequence types (STs) and clonal complexes (CCs) based on the degree of shared alleles at a set of seven house-keeping genes (Dingle et al., 2001). Although ignoring a lot of sequence variation and presence/absence of the accessory genome, MLST has repeatedly shown that certain CCs, such as ST-21 CC and ST-45 CC, are globally distributed in farm animals and are common causes of human infections (Sheppard et al., 2009b; Dearlove et al., 2016). From source attribution studies show that genotypes

predominating in the food animal niche can also be retrieved from wild animals, especially wild birds (Sheppard et al., 2009a; Sheppard et al., 2011). On the other hand, there is growing evidence that in wild birds, *C. jejuni* has strong host association and certain genotypes predominate in specific bird species (Broman et al., 2004; Colles et al., 2008a; Colles et al., 2008b; Sheppard et al., 2011; Griekspoor et al., 2013). Hence, in *C. jejuni*, there are both generalist lineages that can colonize a wide range of host animals and specialist lineages restricted to a few host species, and consequently, specialists and generalists seem to co-exist in many host species including both farm animals and wild birds (Colles et al., 2011; Waldenström and Griekspoor, 2014). Compared to *Salmonella* and *Yersinia* spp. the evolutionary relationship between generalist and specialist lineages of *C. jejuni* is less well understood, as well as the selection pressures behind evolution of specialism or generalism (Sheppard et al., 2014).

Possible explanations for host association of *C. jejuni* genotypes in wild birds could include limited contact between animal species, hence an ecological or behavioural barrier for interspecies transmission. However, there are several examples of wild bird species that share habitat, at least for parts of the year, but still do not seem to exchange *C. jejuni* genotypes (Griekspoor et al., 2013). Other possible factors include differences in diet and feeding behaviour of different bird species, but data indicate that *C. jejuni* genotypes show less association to the host feeding behaviours and more strongly to taxonomy, where related wild bird species tend to more often carry the same, or closely related *C. jejuni* genotypes across large spatial scales (Griekspoor et al., 2013). An alternative explanation would be bacterial adaptation to the intestinal environment of the host, which is likely related to phylogeny. This could include the ability to adhere to and invade intestinal epithelial cells of a particular species, or adaptation to the host immune system and competition with the host's intestinal microbiota. Indeed, there is evidence that specific genera in the host microbiota can reduce colonization resistance to *Campylobacter* (Bereswill et al., 2011; Haag et al., 2012; Dicksved et al., 2014) suggesting that different microbiota composition between species can constitute barriers for transmission. Such adaptations could have evolved through long periods of co-existence and resulted in *C. jejuni* lineages restricted to taxonomically related birds (Waldenström and Griekspoor, 2014).

If limited contact between wild bird species, or differences in diet or feeding behaviour was the reason behind the strong host association, experimental infection of wild birds with *C. jejuni* strains of different origins would probably yield similar colonization patterns. On the other hand, if a *C. jejuni* strain is adapted to the gut of a certain bird species, it would be expected that challenge of a different bird species with that particular strain would result in reduced colonization. Data in support of this was obtained in an infection experiment using the wild European robin (*Erithacus rubecula*) as a host (Waldenström et al., 2010). In this experiment, robins were challenged with two genetically distant *C. jejuni* strains: one strain, isolated from a human patient (ST-48, ST-48 CC) and another strain, isolated from a song thrush (*Turdus philomelos*) (ST-1315, ST-1304 CC). Whereas the song thrush isolate successfully colonized the birds for up to ten days, the human isolate failed to colonize the birds. However, although taxonomically related to the *Turdus* genera with species frequently carrying *Campylobacter* spp., European robins are infrequent carriers of *C. jejuni* in nature (Waldenström et al., 2002).

To determine if *C. jejuni* isolated from one bird species would incur decreased colonization ability in a different bird host, an *in vivo* challenge experiment in a more relevant model species, the mallard (*Anas platyrhynchos*), was used. Mallards have high prevalence of *C. jejuni* in nature, and can carry many different genotypes simultaneously (Colles et al., 2011; Griekspoor et al., 2013; Mohan et al., 2013). *C. jejuni* belonging to many different CCs have been detected in mallards including such commonly found in humans, farm animals and other wild birds (www.pubmlst.org/campylobacter/, 20151229). Colonization in mallards was studied using combinations of single infection and competition experiments with *C. jejuni* strains isolated from three different bird species [song thrush, domestic chicken (*Gallus gallus domesticus*) and mallard]. Genetic relationships between strains were studied by whole genome comparisons both between the three strains and between the pangenomes of the CCs that the strains belonged to. The hypothesis that the *C. jejuni* strains differ in colonization ability, with presumed highest ability in those strains with a known genotypic host association with the model host, was tested. The aim of this chapter was to: (1) Examine the pangenomes of three distinct *C. jejuni* lineages and (2) Identify lineage-specific genes that can be used to monitor bacterial load in the competition animal experiments.

Results

C. jejuni comparative genomics

Phylogenetic analysis of 143 strains including the three strains used in the infection experiments (Figure 5.1, Appendix, Table S5.1), revealed that the genetic distance between the song thrush strain and the mallard strain was 1.5 times greater than that between the chicken strain and the mallard strain. Pangenome comparison of the STs of the three strains used for infection using the 143 strains, identified one ST-specific unique gene out of 1846 in strain #65 (ST-104, ST-21 complex). Additionally, 14 ST-specific genes out of 4,993 genes were found in the three mallard strains examined and 20 ST-specific genes out of 10,746 genes in the six song thrush strains (Appendix, Table S5.2). ST-specific genes were also used as candidates for strain-specific qPCR targets. The specificity of genes *id4678_0651* for the mallard strain, and *id65_1178* for the chicken strain, was confirmed *in vitro* and these targets were subsequently used for the monitoring of strain dynamics during the two competition experimental inoculations of this work.

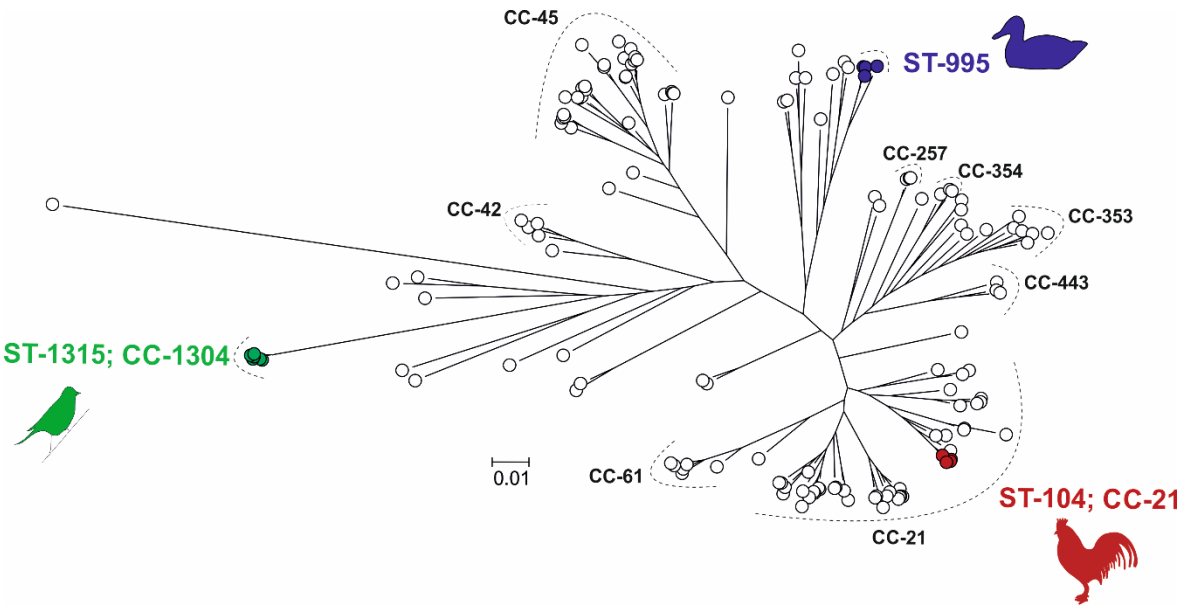


Figure 5.1. Population structure of 142 *Campylobacter jejuni* strains. Coloured *C. jejuni* strains were chosen to represent pan-genomes for the three ST-types used in this chapter, including ST-995 isolates from mallards (blue), ST-1315 (ST-1304 CC) isolates from song thrushes (green) and ST-104 (ST-21 CC) isolates from broiler chickens (red). One strain from each group was selected for experimental infection of mallards. The phylogenetic tree was reconstructed using an approximation of the maximum-likelihood algorithm in RAXML. The scale bar indicates the estimated number of substitutions per site. Blank circles denote *C. jejuni* genomes added to the analysis to provide phylogenetic context to the strains of interest.

Experiment 1: Challenge of mallards with the three *C. jejuni* strains in separate groups.

In experiment 1, each of the three groups of mallards was exposed to one of the three strains from mallard, chicken or song thrush. As shown in Figure 5.2, Figure 5.3, Figure 5.4 and Table S5.3 (Appendix), important differences in the dynamics of bacterial colonization were observed between the groups during the experiment. The birds exposed to the mallard strain excreted high numbers of bacteria [mean 10^4-10^6 colony forming units per millilitre (cfu/ml)] throughout the experiment, 1-18 days post infection (dpi). The birds exposed to the chicken strain had overall lower levels of bacteria in faeces, with peak mean levels of 10^4 cfu/ml. At 18 dpi, only two out of six birds exposed to the chicken strain excreted *C. jejuni*. The song thrush strain was detected at 10^3-10^4 cfu/ml in faeces the first few days after exposure, but bacterial levels declined rapidly. After 7 dpi, the strain could only be detected in two birds and at 18 dpi, the strain was only detected in the caecum of one bird. The mallard strain produced significantly higher bacterial loads, both when analysing all strains together ($Mean_1$, including data from all sampling days, mallard vs chicken vs song thrush, n=30; χ=20.9; df=2; p<0.0001, Kruskal-Wallis test), and by direct comparison between the mallard strain and the chicken strain or the song thrush strain, respectively ($Mean_1$, mallard vs chicken, n=20, p<0.0002 and mallard vs song thrush, n=20; p=0.0003, Mann-Whitney *U* test). By direct comparison between the chicken strain and the song thrush strain, the chicken strain produced significantly higher bacterial loads ($Mean_1$ chicken vs song thrush, n=20; p=0.02104, Mann-Whitney *U* test). No *Campylobacter* spp. was detected in fecal samples from the birds prior to inclusion in the experiments. Control experiments were performed to assess the survival of the three *C. jejuni* strains at room temperature in the water used in the experiments. These revealed a rapid loss of viability and none of the strains survived after 12h. The fractions of each strain surviving in the water after 6h were 0.40% for the mallard strain, 0.46% for the chicken strain and 0 for the song thrush strain (SD=0.52, 0.56, 0). [N.B. The infection experiments and statistics in experiment 1 were performed by Clara Atterby, Uppsala University, Uppsala, Sweden.]

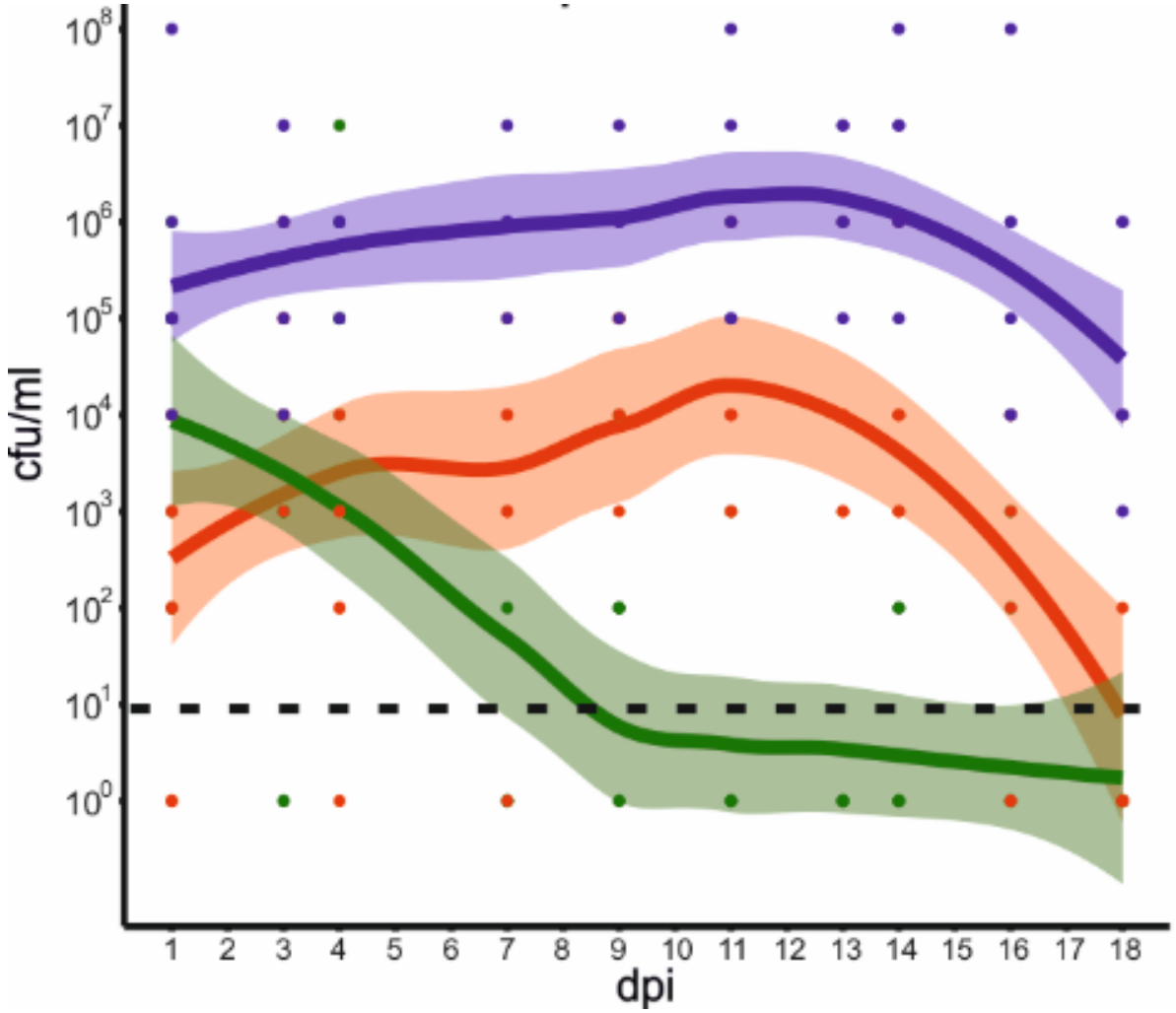


Figure 5.2. **Experimental infection 1 of mallards with *C. jejuni* strains of various hosts.** Colonization dynamics in mallards during infection with *C. jejuni* strains isolated from mallard (purple), chicken (red), and song thrush (green). The graph illustrates the predicted smoothed mean value for each strain with 95% confidence bands based on the mean colony forming units (cfu) per ml of initial suspension for all fecal samples at each time point, as measured by plate counts. Each dot represents fecal cfu/ml from one bird at each time point. Zeros were replaced for one to fit a log scale. The dashed line indicates the theoretical limit of detection.

Experiment 2: Consecutive challenge with the chicken and mallard strain in the same group

In experiment 2, strain specific colonization ability was further assessed by competition experiments where four birds were challenged consecutively with the chicken and the mallard strains. The experiment was performed twice and in experiment 2a, the chicken strain was introduced at day 0 and the mallard strain at 4 dpi. In experiment 2b the two strains were introduced to a new group of birds in the reverse order. In experiment 2a, the chicken strain had established colonization in all four birds at 1 dpi (Figure 5.3). Bacterial numbers started to decrease in 2 birds at 4dpi and were markedly reduced in all birds at 5 dpi, one day after the introduction

of the mallard strain. The chicken strain remained at a level of approximately 10^2 cfu/ml until 9 dpi, but was thereafter no longer detectable in faeces from any of the mallards, except for one bird that again shed high numbers of bacteria at 18 dpi. The mallard strain, introduced at 4 dpi, was detected in high numbers in faeces on day 6, and remained high until day 11 dpi. Thereafter it decreased in abundance but remained at a level of 10^2-10^3 cfu/ml until the end of the experiment (Figure 5.3).

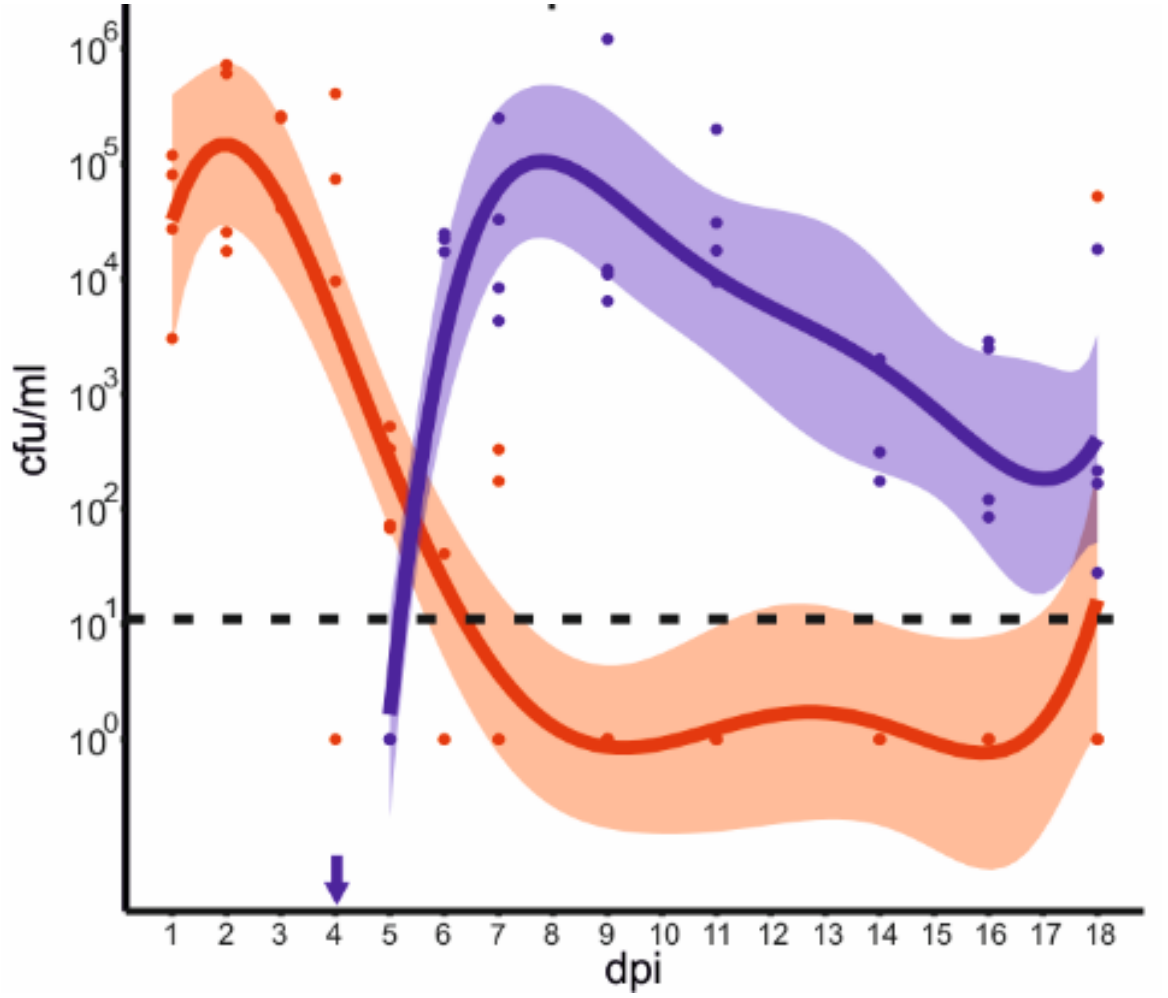


Figure 5.3. **Experimental infection 2a of mallards with *C. jejuni* strains of various hosts.** Colonization dynamics in mallards during mixed infection with strains isolated from mallard (purple) and chicken (red). The graphs illustrate the predicted smoothed mean value for each strain with 95% confidence bands corresponding to the mean cfu/ml of initial suspension for all fecal samples at each time point, as determined by real-time PCR with CT-values transformed to cfu/ml. In experiment 2a, birds were infected at 0 dpi with the chicken strain followed by the mallard strain at day 4 dpi (indicated by blue arrow). Zeros were replaced for 1 to fit a log scale. The dashed line indicates the theoretical limit of detection.

In experiment 2b, the mallard strain had established colonization of all birds at 1 dpi. In contrast to the chicken strain in experiment 2a, there was no decrease in numbers

of the mallard strain at 5 dpi and on. Instead, an increase in bacterial numbers of this strain was observed (Figure 5.4, Appendix, Table S5.4). The mallard strain remained in high abundance in the fecal samples throughout the experiment. The chicken strain, introduced at 4 dpi, could be detected in low numbers in faeces 5 dpi. However, bacterial numbers in fecal samples did not peak until 7 days after inoculation (11 dpi).

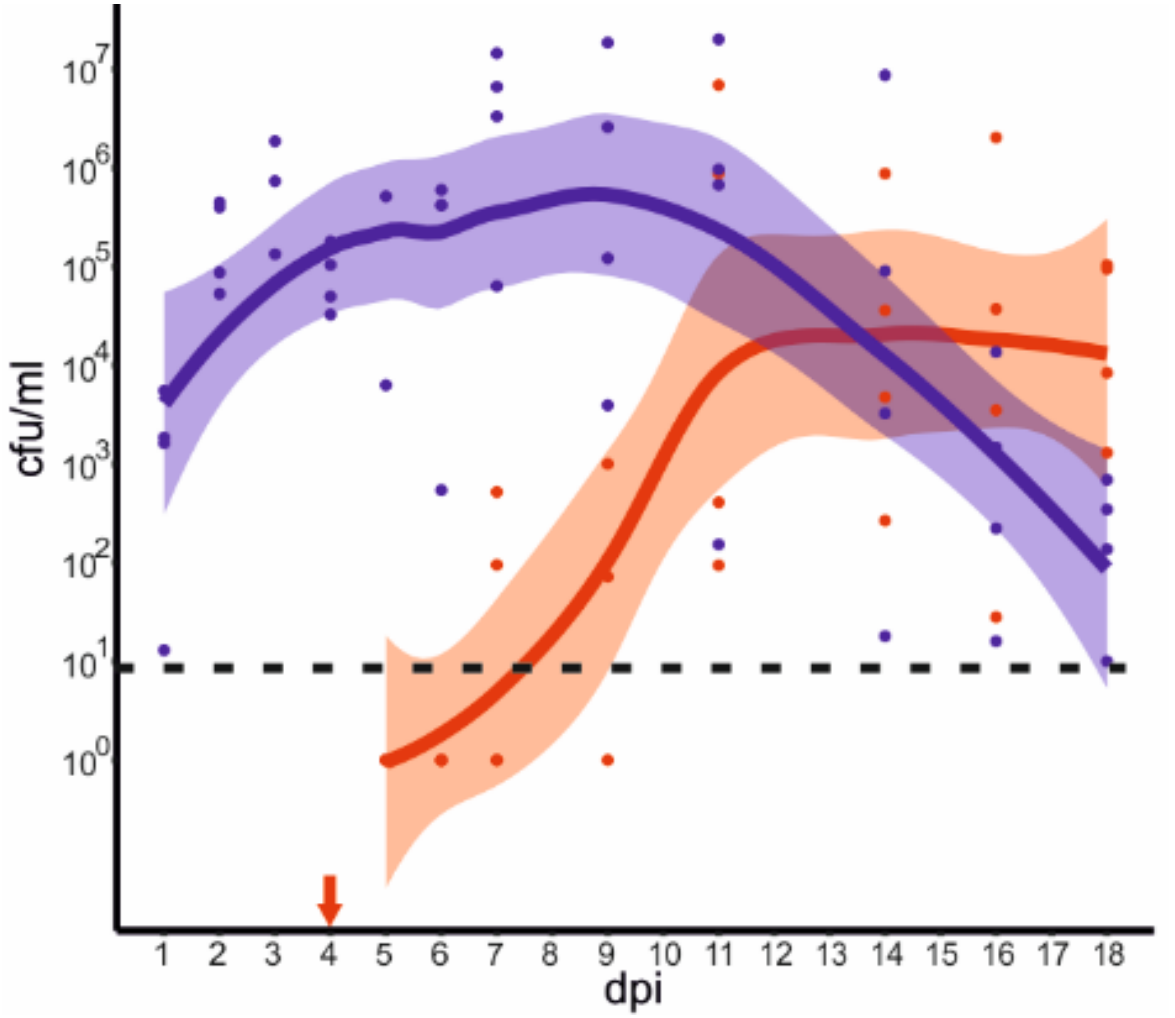


Figure 5.4. **Experimental infection 2b of mallards with *C. jejuni* strains of various hosts.** Colonization dynamics in mallards during mixed infection with strains isolated from mallard (purple) and chicken (red). The graphs illustrate the predicted smoothed mean value for each strain with 95% confidence bands corresponding to the mean cfu/ml of initial suspension for all fecal samples at each time point, as determined by real-time PCR with CT-values transformed to cfu/ml. In experiment 2b, birds were infected at 0 dpi with the mallard strain followed by the chicken strain at 4 dpi (indicated by red arrow). Zeros were replaced for 1 to fit a log scale. The dashed line indicates the theoretical limit of detection.

In comparisons between experiments, the chicken strain produced significantly higher bacterial loads the first days after introduction in experiment 2a compared to bacterial loads of the same strain the first days after introduction in experiment 2b,

when the mallard strain was already present (Chicken 2a $Mean_A$, 1-4 dpi vs Chicken 2b $Mean_B$, 5-9 dpi, n=8, p=0.0286, Mann-Whitney U test). The corresponding comparison for the mallard strain yielded no significant differences (Mallard 2b $Mean_A$, 1-4 dpi, vs Mallard 2a $Mean_B$, n=8; p=0.4857, Mann-Whitney U test). Comparisons of faecal bacterial loads between the two strains on the first four days when introduced as the second strain, revealed significantly higher bacterial loads of the mallard strain compared to the chicken strain (Mallard 2a $Mean_B$, vs Chicken 2b $Mean_B$, p=0.0286, Mann-Whitney U test). [N.B. The infection experiments and statistics in experiment 2 were performed by Evangelos Mourkas, Uppsala University, Uppsala, Sweden.]

Bacterial colonization at different sites in the mallard gastrointestinal tract

In experiment 1, two birds were sacrificed in each group on 1 and 3 dpi for assessment of bacterial loads at different sites along the gastrointestinal tract. At these time points, all strains were found in high numbers in the gizzard, jejunum, caecum and colon. On 18 dpi, all remaining birds were sacrificed, and bacterial counts were assessed at the same sites of the GI tract. The mallard strain was detected in all GI-segments with high bacterial loads (10^6-10^8 cfu in caecum, 10^3-10^5 cfu/ml in colon, 10^2 cfu/ml in gizzard and 10^3 cfu/ml in jejunum) (Table 5.1).

Table 5.1. Oligonucleotide primers designed for real-time PCR assay.

Gene name	*C. jejuni* strain	Designation	Sequence	PCR product size	*T*m (°C)
id65_1178	Chicken strain	Fwd-Ch-hypr 241	5'-GTCGTACAGGATTTT ATGATGAGAG – 3'	241	61.5
id65_1178	Chicken strain	Rev-Ch-hypr 241	5'-CGGCAACTTTTATAA TCAGCTT – 3'	241	60.3
id4678_0651	Mallard strain	Fwd-Mal-unch 209	5'-CAATCGCCTCTTAAA TCTCCA – 3'	209	60.7
id4678_0651	Mallard strain	Rev-Mal-unch 209	5'-AAATCTGAATGCGGT GGAAG – 3'	209	61.4

In contrast, the chicken strain could not be detected at all in gizzard or jejunum and was found in moderate numbers (10^2-10^3 cfu/ml) in five out of six birds in the caecum and/or colon. Although negative on fecal sampling, three birds carried the chicken strain in the caecum at 18 dpi. The song thrush strain was only detected in the

caecum of one bird, and this bird was negative on fecal sampling. [N.B. The necropsies and data analysis in experiment 1 were performed by Clara Atterby, Uppsala University, Uppsala, Sweden.] In experiment 2a and 2b, the mallard strain and chicken strain showed similar bacterial loads in the caecum and colon 18 dpi, but only the mallard strain was detected in jejunum and gizzard (Table 5.1). Caecum was the preferred site of colonization in both experiment 1 and 2 for all strains (Table 5.1). In experiment 1, the mallard strain produced significantly higher bacterial loads in caecum compared to all other sites (mallard caecum vs mallard gizzard, p=0.0001, mallard caecum vs mallard jejunum, p=0.0001, mallard caecum vs mallard colon, p=0.0023, Mann-Whitney *U* test). The chicken strain produced significantly higher bacterial loads in caecum compared to gizzard and jejunum (chicken caecum vs chicken gizzard, p=0.0007, chicken caecum vs chicken jejunum, p=0.0007, Mann-Whitney *U* test) but there was no significant difference between bacterial loads in cecum and colon (chicken caecum vs chicken colon, p=0.2310, Mann-Whitney *U* test). For the song thrush strain, there was no significant differences in bacterial loads between the different sites of the GI tract (song thrush caecum vs song thrush gizzard, p=0.2310, song thrush caecum vs song thrush jejunum, p=0.0759, song thrush caecum vs song thrush colon, p=0.5599, Mann-Whitney *U* test). [N.B. The necropsies and data analysis in experiment 2 were performed by Evangelos Mourkas, Uppsala University, Uppsala, Sweden.]

General health parameters

All ducks behaved normally, and no clinical symptoms of disease were observed in any of the experiments. A slight decrease of body mass was observed in all three experiments during the first five days, but after 4 dpi the weight remained stable until the end of the experiments. A similar decrease in body mass was observed also in uninfected mallards when they were moved from the bigger flock into the experimental rooms and is likely due to handling and settling stress in a new environment. No macroscopic evidence of inflammation or lesions was observed in any of the internal organs during necropsy.

Discussion

Zoonoses account for the majority of human diseases and many zoonotic pathogens are transmitted from wild to domestic animals and further to humans. Understanding the underlying factors and host-pathogen interactions that determine the outcome

of interspecies transmission events for zoonotic, multi-host pathogens is important for adequate planning of interventions to reduce spread to farm animals and ultimately, to protect humans from infection. In this chapter, barriers for transmission of *C. jejuni* between bird species, were assessed. Such knowledge can increase understanding of the spread of this pathogen from its natural source to domestic animals and further to humans.

Epidemiological evidence suggests very limited interspecies transmission of specialist *C. jejuni* lineages between wild birds, but less is known about the underlying factors behind this observation (Broman et al., 2004; Colles et al., 2008a; Colles et al., 2008b; Sheppard et al., 2011; Griekspoor et al., 2013). The hypothesis that reduced colonization ability due to adaptation to a different host species could account for this pattern was tested by assessing differences in interspecies colonization ability between a generalist and a specialist *C. jejuni* strain isolated from different bird species using captive wild mallards as model host. In single infection experiments, clear differences were observed in colonization ability between the strains, consistent with expectations from their genetic backgrounds. The mallard strain was the best colonizer with the highest amounts of bacteria excreted in the faeces throughout the experiment. In contrast, the song thrush isolate showed significantly reduced colonization ability and was only detectable in 2 out of 6 birds after 7 dpi. The chicken strain was secreted throughout the experiment but in significantly lower numbers compared to the mallard strain. These strain specific differences in colonization ability were further emphasized by their relative abundance observed in different parts of the gastrointestinal tract, where the mallard strain could be detected in several parts of the intestine, whereas the two other strains mainly were restricted to caecum and colon.

Although the differences in colonization ability appeared smaller between the chicken- and the mallard strain compared to that between the mallard and the song thrush strain in single infection experiments, the results of the competition experiments revealed a clear competitive advantage of the mallard strain compared to the chicken strain, based on a number of observations. The mallard strain when introduced on day 4 dpi, rapidly outcompeted the chicken strain with peak shedding already 2-3 days after inoculation (Figure 5.3), whereas the chicken strain when introduced on day 4 dpi, needed more time to reach peak shedding (Figure 5.4).

The chicken strain dropped sharply in numbers shortly after the mallard strain was introduced (Figure 5.3), but no corresponding decrease of the mallard strain was seen in response to introduction of the chicken strain (Figure 5.4). As a whole, faecal shedding of the chicken strain seemed negatively affected by competition with the mallard strain, whereas no significant effect was seen for the mallard strain. Instead, this strain colonized significantly better than the chicken strain under competition, as determined by the relative shedding of the two strains when both were introduced as the second. Some variation in the level of colonization was observed for the same strains between experiment 1 and 2 and these differences were likely due to the fact that colonization was monitored by PCR in experiment 2 as well as differences in the age and number of birds between these experiments. Control experiments assessing the survival of the strains in water suggest that all strains were short lived under the experimental conditions. Although the song thrush strain survived for a shorter time period compared to the other two strains, this did not seem to have any impact on the establishment of initial colonization as mallards inoculated with that strain shed more bacteria 1dpi compared to those inoculated with the chicken strain. Furthermore, as the water pools and the experiment rooms were cleaned every 24h, differences in long term bacterial survival in the environment are not likely to have had a large impact on the observed colonization patterns in the birds.

Taken together, these results show clear strain specific differences in the ability to colonize the mallard gastrointestinal tract. The differences correspond well with the phylogenetic relatedness of the strains and are likely associated with host origin. Genomic analysis revealed a greater genetic distance between the mallard strain and the song thrush strain compared to that between the mallard and the chicken strain. This pattern corroborates what was even more clearly seen in an earlier study using 2,294 *C. jejuni* strains from wild birds, domestic chickens and humans (Griekspoor et al., 2013). Although the number of song thrush and mallard isolates used in the present work was small, comparison of pangenomes from the CCs of each of the three strains revealed important differences in their gene content. Some of these genes could possibly explain the observed differences in the ability to colonize the mallard gastrointestinal tract. However, an accurate analysis of such gene ontology would require more strains from song thrushes and mallards as well as a rigorous panel of *in vitro* assays. At this point, speculations around possible factors making up the barriers for colonization of the mallard (or any bird's) intestine

and the host specific *C. jejuni* genes that are needed to overcome them, can be made. The host's immune response is always an important factor in infection and both innate- and humoral immunity is likely involved. However, although the immune response was a likely cause of the reduction in bacterial numbers observed towards the end of the experiments, it is less likely that specific immunity accounted for the dramatic effects seen in the competition experiments or in the rapid reduction of bacterial numbers of the song thrush strain. Instead, results suggest that other factors related to the intestinal environment of specific bird species might make up such barriers. These could include the intestinal microbiota, structure and glycosylation of the mucin layer, structure of receptors expressed at the epithelial surface or other factors that would favour strains that have co-evolved with its host. Such barriers would lead to lower bacterial load, less shed bacteria and hence, fewer potential transmission events of the new strain. In other words, a reduced fitness of the strain in the population, especially in competition with other strains that are better adapted to the host. Hence, in the absence of competing strains, a less-than-optimally adapted strain may still successfully transmit, but given the competitive landscape of *C. jejuni* in birds, the chance of long-term proliferation is reduced. Consistent with this, it may be expected that a generalist *C. jejuni* genotype would have better colonization ability in a new host compared to a specialist genotype adapted to a different host species, as was the case in this work.

This work was performed in captive wild mallards, and the results cannot be directly extrapolated to infection of chickens as there are important differences between these species in terms of anatomy, food intake, behaviour etc. Attempts to study differences in colonization ability between *C. jejuni* strains in chickens have been made through infection experiments (Glunder, 1995; Korolik et al., 1998; Stas et al., 1999; Hänel et al., 2004; McCrea et al., 2006; de Haan et al., 2010; Chaloner et al., 2014). However, these studies have mainly focused on *C. jejuni* isolates from farm animals and humans and although results are somewhat conflicting between studies, they collectively suggest that host origin is a less important determinant of colonization ability in chickens when comparing *C. jejuni* isolates from such sources. This can likely be explained by the fact that most farm animals and humans share *C. jejuni* strains with similar or identical genotypes and that intensively reared chickens are susceptible to multiple genotypes of *C. jejuni* (Frances M. Colles et al., 2008; de Haan et al., 2010; Griekspoor et al., 2013; Griekspoor et al., 2015). On the

other hand, an epidemiological study assessing natural transmission of wild bird associated *C. jejuni* strains to free range broiler chickens, suggested limited or no transmission between these bird species despite the fact that the birds occupied the same ranging area. Infection experiments in chickens with *C. jejuni* strains of wild bird origin should be performed to assess the risk of spread of such strains to broilers. This is important as contact with broiler chickens or broiler meat is the most common source of human *Campylobacter* infection. If all *C. jejuni* strains can be readily transmitted to broiler chickens, then the wild bird population constitutes an infinite source of new strains that can feed into the chicken population with increased risk of transmission to humans. Direct transmission of specialist *C. jejuni* lineages from wild birds to humans is rare (Griekspoor et al., 2013), and although the results from the mallard infection experiments in this chapter cannot be extrapolated to infection in humans, it is possible that reduced ability of such strains to colonize the human gut could be the reason behind this.

In conclusion, this work shows that experimental infection with *C. jejuni* strains in a natural host results in different colonization outcomes depending on the host origin of the strain. These results suggest that the strong host association observed in *C. jejuni* from wild birds is likely not due to the absence of direct or indirect contact between these host species. Instead, the barriers for interspecies transmission may be more directly associated to strain specific differences in colonization ability that are likely related to host origin of the bacterial strain as well as to physiological factors of the host.

Materials and methods

Bacterial strains and genomes

A total of nine strains were isolated and genome sequenced as part of this chapter. Three *C. jejuni* strains of different host origin were used in the two bird infection experiments. These included strain #3927 (ST-995) (unassigned CC), isolated from a mallard in Sweden in 2002. This ST-type has previously been isolated from chickens and dogs (www.pubmlst.org/campylobacter/, 20151229). Strain #3926 (ST-1315 in ST-1304 CC) was isolated from a song thrush captured in Sweden in 2000. This strain is the only reported strain of this ST-type and belongs to a CC that appears restricted to thrushes. Strain #65 (ST-104, in ST-21 CC), was isolated from a broiler chicken in the UK 2006. Strains of ST-104 have been frequently isolated

from poultry, humans, several other animal species and from the environment (www.pubmlst.org/campylobacter/, 20151229). Other strains in this CC have also been found in gulls (www.pubmlst.org/campylobacter/, 20170629). Strains were sampled from song thrushes and mallards between 2000 and 2002. (Appendix, Table S5.1). Apart from the strain from chicken isolated prior to this chapter work in the UK, all strains were isolated from wild birds captured at the Ottenby Bird Observatory, Öland, Sweden. A total of 134 *C. jejuni* genomes from two previously published studies (Sheppard et al., 2014; Sheppard et al., 2013) (Appendix, Table S5.1) were added to the dataset of this chapter to provide a population-wide phylogenetic context for the strains used for infection experiments, as well as to identify genes that are specific to them but not found in a broader population.

The three bacterial strains used for inoculation were minimally passaged on agar plates during isolation and stored at -80ºC until used. Bacterial inocula were prepared from frozen stocks by culture on blood agar plates (Columbia agar II containing 8% [vol/vol] whole horse blood) under microaerobic conditions using GENbox anaer (Biomerieux, Askim, Sweden) with CampyGen 2.5L Atmosphere Generation Systems Packs (Oxoid, Basingstoke, UK) for 24 to 48h at 42°C. Bacteria were harvested and suspended in PBS (pH 7.4). Optical densities were measured using a UVmini-1240 UV-Vis Spectrophotometer (SHIMADZU, Lidingö, Sweden) and cell densities were adjusted to 1×10^9 colony forming units (cfu)/ml. The bacterial concentrations of the inocula were verified by culture on blood agar plates and were all within the range of $0.6–1.8\times10^9$ cfu/ml.

DNA extraction, genome sequencing, assembly and archiving

The genome sequences of nine strains were obtained. Briefly, DNA was extracted using the QIAmp DNA Mini Kit (Qiagen, Crawley, UK) according to the manufacturer's instructions. Nucleic acid content was quantified on a Nanodrop spectrophotometer prior to normalization and sequencing. Nextera XT (v3 technology, 250bp paired-end) libraries were prepared and high-throughput sequencing was performed using an Illumina MiSeq benchtop sequencer (Illumina, San Diego, CA). Short reads were assembled de novo using SPAdes (version 3.0.0) and evaluated using QUAST (Bankevich et al., 2012; Gurevich et al., 2013). Assembled DNA sequences were uploaded to a web-based database based on the BIGSdb platform (Jolley et al., 2018) which allowed the archiving, whole genome

gene-by-gene sequence alignments and prevalence analyses. Novel *C. jejuni* genome sequences are available publicly online (NCBI BioProject: PRJNA415188).

Reference pan-genome, phylogenetic tree inference and qPCR targets

A reference pan-genome list was assembled using a previously published method (Méric et al., 2014) from the whole genomes of 13 isolates. Six of the nine newly sequenced strains were isolated from song thrushes (ST-1304 CC) and the remaining 3 from mallards (ST-995 CC). These were augmented with four genomes isolated from chicken, cattle and human (ST-21 CC) in order to get a representative number of isolates for each CC used in the pan-genome (Appendix, Table S5.1). Briefly, automatic annotations were obtained using RAST (Aziz et al., 2008) and from a total of 22,060 genes detected, 2,489 genes were present in all isolates after the removal of allelic variants using BLAST, with alleles of the same gene being defined as sequenced of >70% sequence identity on >10% of the sequence length (Parkhill et al., 2000; Méric et al., 2014). A whole-genome multiple sequence alignment was obtained by gene-by-gene orthologue identification using MAFFT (Katoh, 2002), and concatenation into a single contiguous sequence for input and phylogenetic tree reconstruction using the approximation of the maximum-likelihood algorithm implemented in RAxML (Stamatakis, 2006), running on CLIMB cloud-computing servers (Connor et al., 2016).

Prevalence and allelic variation for every gene of the reference pan-genome list in 143 *C. jejuni* genomes (Appendix, Table S5.1) was determined using BLAST, as previously published (Méric et al., 2014; Méric et al., 2015; Pascoe et al., 2015; Yahara et al., 2017). Specifically, genes found in the strains that were used for infection of the birds, and absent in other strains, were considered as candidate targets for development of primers for a quantitative real-time PCR, targeting specifically each of the two strains used in the competition infection experiment. Primers were designed using the online “Primer 3 input software version 0.4.0” (Koressaar and Remm, 2007; Untergasser et al., 2012). After evaluation of several primer candidates two primer pairs were selected (Table 5.2). The specificity of all primers was assessed by BLAST in the Genbank public repository and evaluated by analysis of fecal samples from *Campylobacter* negative mallards. [N.B. The primer selection analysis was performed by Evangelos Mourkas, Uppsala University, Uppsala, Sweden.]

Table 5.2. Sampling days and statistical analysis for experiment 1 and 2*.

Experiment	Mean	Sampling days											
Exp 1	$Mean_1$	1	[a]	3	4			7	9	11	14	16	18
Exp 1	$Mean_2$	1		3	4			7	9	11	14	16	18
Exp 2	$Mean_B$	1	2	3	4	5	6	7	9	11	14	16	18
Exp 2	$Mean_C$	1	2	3	4	5	6	7	9	11	14	16	18

*All sampling days are indicated as numbers in the table. Shaded boxes display from which days data were included in each mean value used for statistical analysis. For experiment 1, two mean values were calculated from the fecal cfu count of each of the 30 birds. $Mean_1$ was based on data from all sampling days and $mean_2$ was based on data from week two and three. For experiment 2, two mean values were calculated from the fecal cfu counts of each of the 8 birds. $Mean_B$ was based on data from the first four days after introduction of the first *C. jejuni* strain in experiment 2a and 2b respectively, and $mean_C$ was based on data from the first four days after introduction of the second *C. jejuni* strain.

[a] No samples were obtained on day 2, 5 and 6 in experiment 1.

Mallard infection model and housing

The mallard infection model has been used for studies of influenza A virus, and has been described in detail previously (Järhult et al., 2011). Briefly, one day post hatch, male mallards were introduced to the biosecurity level two (BSL2) animal facility at the Swedish National Veterinary Institute (SVA). The mallards were housed indoors with access to pools for swimming and feed and water *ad libitum*. The experimental rooms were HEPA filtered with positive air pressure and double doors and held one pool with water each. Strict hygiene regulations were followed by the staff when handling the mallards and moving between rooms. Before inclusion in the experiments, all birds were tested negative for fecal *Campylobacter* spp. growth on modified charcoal cefoperazone agar (mCCDA) plates (Department of Clinical Microbiology, Uppsala University Hospital).

Two different experimental setups were applied to study the colonization ability of the *C. jejuni* strains (Figure 5.5). In experiment 1, 10 mallards (8 weeks of age) were placed in each of three separated experimental rooms. Each group was exposed to one of the three *C. jejuni* strains, mallard (#3927), chicken (#65) and song thrush

(#3926) (Appendix, Table S5.1), on day 0. Fecal content was obtained from all birds in experiment 1 on 0, 1, 3, 4, 7, 9, 11, 14, 16 and 18 dpi. [N.B. The infection experiment 1 was performed by Clara Atterby, Uppsala University, Uppsala, Sweden.] Experiment 2 was designed to study how the colonization ability of the *C. jejuni* strains of chicken origin (#65) and mallard origin (#3927) was affected by competition with each other within the same bird. Groups of four birds were consecutively infected with the two strains at different time points and the experiment was repeated introducing the strains in the reverse order. In experiment 2a, the birds (24 weeks of age) were exposed to the chicken strain day 0 and to the mallard strain 4 dpi. In experiment 2b, birds (27 weeks of age) were exposed to the mallard strain day 0 and to the chicken strain 4 dpi. Fecal samples were obtained from all birds in experiment 2 on 0, 1, 2, 3, 4, 5, 6, 7, 9, 11, 14, 16 and 18 dpi. In all experiments, exposure was obtained by adding bacterial inoculum to the water pool yielding a bacterial concentration of approximately 5×10^4 cfu/ml of water. The water used was non-chlorinated tap water. During the exposure days, the water pool used for swimming was the only source of drinking water in the experiment room. The pool was emptied and thoroughly rinsed with fresh water every 24 hours, including after inoculation. This route of exposure was chosen to simulate a natural situation where birds get infected from contaminated water in their environment. The high bacterial concentration was chosen to make sure that all ducks would ingest viable bacteria. Control experiments were performed to assess the survival of the three strains in water from the same source as used in the animal experiments and at the same temperature (room temperature). Bacterial inocula were prepared as described above and added to 100ml of water in Erlenmeyer flasks to yield a concentration of appr. 10^4 cfu/ml. Subsamples of 100µl were withdrawn at time 0, 6h, 12h and 24h after inoculation and bacterial numbers were assessed by culture on mCCDA plates. The experiment was performed twice with triplicate flasks.

All birds were tagged with color-coded rings for identification. Each bird was visually examined daily for gross signs of injury or lethargy and was subsequently placed in a clean cardboard box where it was weighed and left to defecate. Faeces was collected from each box using a sterile cotton swab (ClassiqSwabs, Copan, flock technologies, Täby, Sweden). In experiment 1, two birds from each experimental room were sacrificed on 1 and 3 dpi, and the remaining birds in both experiments were sacrificed on day 18 dpi by euthanasia with pentobarbital injected in the tarsal

vein. Necropsies were performed after euthanasia and samples of intestinal contents from the gizzard, jejunum (distal to Meckel's diverticulum), caecum and colon were obtained. All samples were stored on ice in Luria Broth containing 20% glycerol and analysed for bacterial growth within 4 hours. [N.B. The infection experiment 1 was performed by Clara Atterby, Uppsala University, Uppsala, Sweden.]

Bacterial quantification in fecal samples

In experiment 1, plate counts were enough to follow the colonization in the three groups as each group was exposed to only one strain of *C. jejuni*. [N.B. Plate counts for experiment 1 were performed by Clara Atterby, Uppsala University, Uppsala, Sweden.] In experiment 2, it was impossible to separate the two strains through phenotypical appearance on the plate and we therefore developed a specific real-time-PCR assay. For bacterial enumeration on agar plates, 100mg of fecal sample was put into 1ml of Luria Broth (LB) supplemented with 20% glycerol (to enable freezing of the sample after bacterial enumeration).

The sample was vortexed and centrifuged at 400 x g for 1 min to pellet gross fecal material. Samples were diluted in ten-fold dilution series in PBS and 100µl from each dilution was plated onto mCCDA plates. The plates were incubated microaerobically at 42°C for 48h before colonies were enumerated. Bacterial concentration in fecal samples was expressed as cfu/ml of the initial suspension (in LB glycerol). Due to the large number of birds in experiment 1, bacterial numbers on plates were estimated to the nearest 10 or 100. [N.B. Plate counts and dilution series for experiment 2 were performed by Evangelos Mourkas, Uppsala University, Uppsala, Sweden.]

Development of real-time-PCR for bacterial quantification

Mallard and chicken strains were identified with real-time PCR using the SsoAdvanced Universal SYBR Green Supermix (Bio Rad Laboratories AB, Sundbyberg, Sweden) on a CFX96 Optics Module C1000 Thermal Cycler (Bio Rad Laboratories AB, Sundbyberg, Sweden). The reaction mixture consisted of 1 x SYBR Green, 0.3 µM of each of the primers (Table 5.2), 1 µL of template solution and DNase/RNase-free distilled water (Thermo Fischer Scientific, Waltham, MA, U.S.A.) to a final volume of 20 µL. Final cycling conditions were 98 °C for 3 min

followed by 40 cycles of 98 °C for 15 s and 63 °C for 60 s, followed by a dissociation curve ranging from 65-95 °C. All PCR reactions were performed in triplicates.

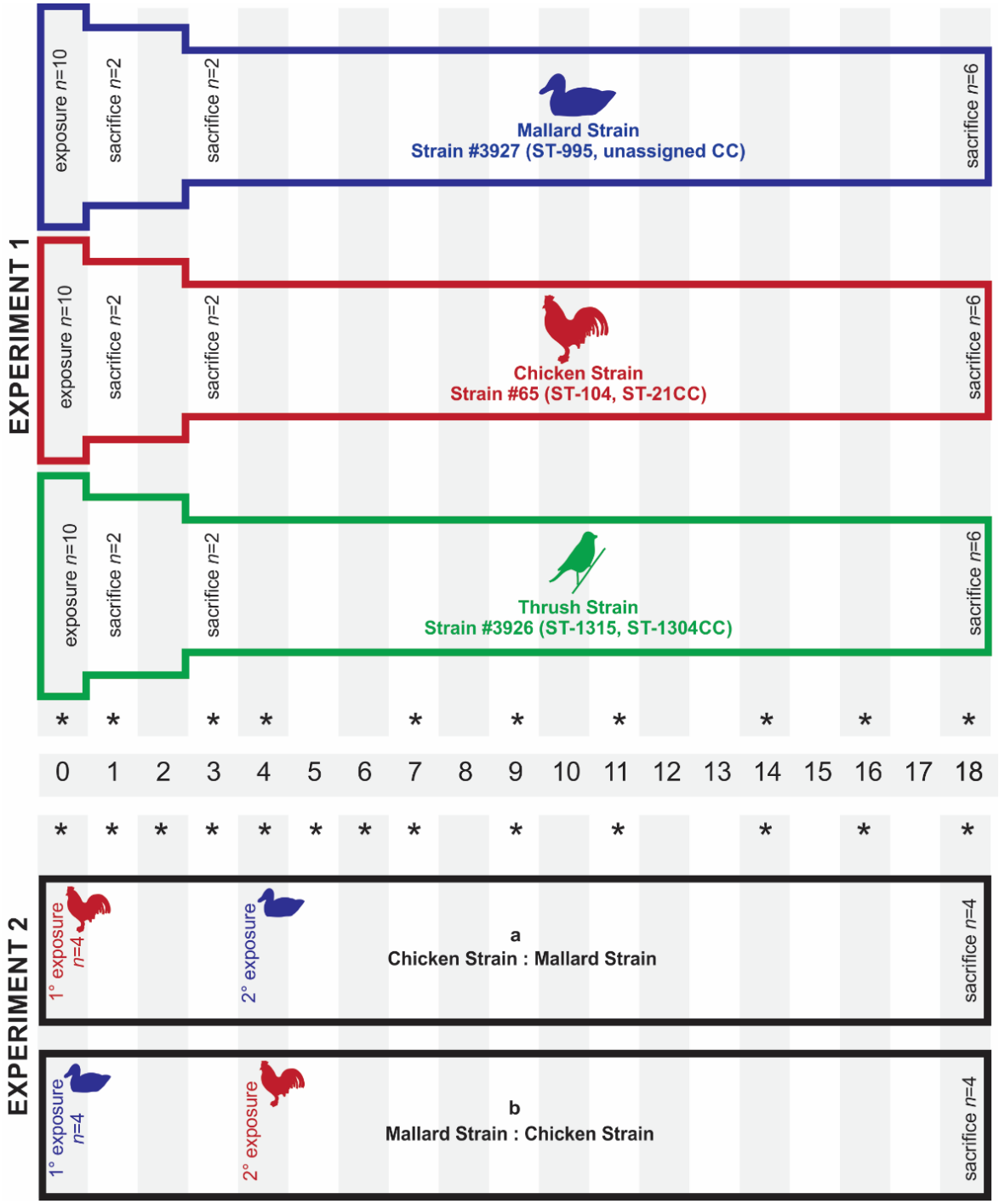


Figure 5.5. Experimental set-up for mallards (*Anas platyrhynchos*) infected with *C. jejuni* of different host origin. Each bar represents a group of birds infected with *C. jejuni*. In experiment 1, each group of mallards was exposed to *C. jejuni* of different host origin (mallard, chicken, song thrush) on day 0. In experiment 2, each group was exposed to one *C. jejuni* strain on day 0 and a second *C. jejuni* strain on day 4. In experiment 1, two birds in each group were sacrificed on day 1 and day 3 respectively, while all remaining birds were sacrificed at the end of both experiments (day 18). The stars indicate the sampling days.

The Bio-Rad CFX Manager 3.1 software (Bio Rad Laboratories AB, Sundbyberg, Sweden) was used for data analyses. The melting point for each amplicon was identified and set as a measure of the specificity of the assay. Primers were designed as described above, and appropriate annealing temperature was assessed using a thermal gradient during optimization. Expected size of the PCR products were verified by agarose gel electrophoresis with a 1.5 % Tris-Acetate-EDTA (Sigma-Aldrich AB, Stockholm, Sweden) agarose gel viewed under UV light together with a GeneRuler™ 100 bp Plus DNA Ladder (Thermo Fischer Scientific, Waltham, MA, U.S.A.).

The C_T-values from the real-time PCR were transformed into cfu using standard curves prepared from each strain (Appendix, Table S5.5). The mallard and the chicken strain were inoculated into Brucella broth and incubated microaerobically for 24h at 42 °C. The concentrations of the two bacterial stocks were quantified on a Nano Drop 2000c spectrophotometer (Thermo Fisher Scientific, Waltham, MA, U.S.A.) and confirmed by plate counts on blood agar plates. An amount of 1.4×10^9 cfu and 2.6×10^9 cfu was used for DNA extraction from the mallard and chicken strain respectively. The extracted DNA was serially diluted to generate a standard curve for each strain. The dilution series was included in each 96 well plate that was analysed in order to allow bacterial quantification as well as to determine the detection limit of the assay. The software constructed slopes of standard curves by linear regression analysis in order to monitor the amplification efficiency and detection sensitivity of every run.

The QIAamp cador Pathogen Mini kit (Qiagen AB, Sollentuna, Sweden) was used for extraction of DNA from the fecal samples. Extraction was performed according to the manufacturer's instructions, with some slight modifications. Briefly, fecal samples were thawed on ice for approximately one hour, vortexed thoroughly for 1 min to ensure homogeneity, and centrifuged at 500 x g for 1 min to pellet gross fecal material. Three hundred microliters of supernatant were drawn and mixed with 200 mg 0.1 mm silica beads cat. no. 11079101z (BioSpec Products, Bartlesville, OK, U.S.A.) and 800 µl of ASL stool lysis buffer (Qiagen AB, Sollentuna, Sweden). Samples were vortexed briefly, incubated in a heating block at 95 °C for 5 min and instantly put on ice for 10 min. This was followed by bead beating in a Bio 101 FastPrep FP120-120V disrupter homogenizer (Savant, Illkirch-Graffenstaden,

France) for 3 x 20 seconds at 5000 rpm, with incubation for 1 min on ice between each cycle. The tubes were then centrifuged at 2500 x g for 1 min to precipitate beads and solid material and 200 µl of the supernatant was used for further extraction according to the manufacturer's instructions.

There was generally a good correlation between the estimates of bacterial numbers by plate counts and real-time PCR over time (Appendix, Table S5.3, Table S5.4). However, the PCR analysis consistently detected one log higher bacterial numbers compared to plate counts. This over estimation was most likely due to the fact that the PCR analysis detected both viable and dead bacteria, in contrast to plate culture. The general level of colonization was lower in experiment 2a compared to 2b as determined both by plate counts and PCR (Figure 5.3, Figure 5.4, Appendix, Table S5.3, Table S5.4). Although some variation was seen between the ducks, they displayed roughly the same colonization pattern. One duck in experiment 2a did not defecate at 6 dpi and 14 dpi while two ducks in experiment 2b had insufficient amount of faeces at 16 dpi and 18 dpi, respectively. [N.B. The above section of analysis was performed by Evangelos Mourkas, Uppsala University, Uppsala, Sweden.]

Statistical analysis

Graphs were generated with ggplot2 package using the loess smoothing function for R software (Wickham, 2009) and illustrate the predicted smoothed mean value for each strain with 95% confidence bands based on the mean cfu/ml of all fecal samples at each time point, as measured by plate counts (Figure 5.2, Figure 5.3, Figure 5.4). The cfu/ml of each bird at each time point is indicated by dots in the graph. Apart from a few time points, every bird had a *C. jejuni* cfu/ml count for each sampling day. The mean cfu/ml counts over the course of several days were calculated for each bird to evaluate overall colonization and colonization at the first days after strain exposure (Table 5.3).

The mean value from each individual bird was grouped with mean values from birds in the same experiment exposed to the same strain. For experiment 1, this resulted in three groups, mallard, chicken and song thrush (n=10/group), and two mean values per group, $mean_1$ and $mean_2$. For experiment 2, this resulted in two groups, mallard and chicken (n=4/group), and two mean values per group, $mean_B$ and

$mean_C$. The groups were compared using nonparametric Kruskal-Wallis one-way analysis of variance and nonparametric Mann-Whitney *U* test.

Table 5.3. Detection of the *C. jejuni* strains in different segments of the gastrointestinal tract of infected mallards.

	Experiment 1									Experiment 2a		Experiment 2b	
Strain origin	Mallard			Chicken			Song Thrush			Mallard	Chicken	Mallard	Chicken
Day (dpi)	1	3	18	1	3	18	1	3	18	18	18	18	18
Gizzard	0/2[a,b]	1/2	1/6	1/2	0/2	0/6	0/2	0/2	0/6	3/4	0/4	2/4	0/4
Jejunum	1/2	2/2	2/6	0/2	1/2	0/6	1/2	0/2	0/6	2/4	0/4	1/4	0/4
Caecum	2/2	2/2	6/6	2/2	2/2	5/6	2/2	2/2	1/6	4/4	3/4	4/4	4/4
Colon	2/2	2/2	6/6	2/2	2/2	2/6	2/2	2/2	0/6	4/4	1/4	4/4	3/4

[a] Number of mallards in which *C. jejuni* was detected in the indicated segment of the gastrointestinal tract out of the total number of infected mallards investigated at each time point.

[b] The theoretical limit of detection of the assay is 10^1 cfu/ml.

For the necropsy cfu counts in experiment 1, every bird had a *C. jejuni* cfu/ml count for each organ (gizzard, jejunum, colon and caecum). The cfu/ml counts from all birds exposed to the same strain were grouped in the respective organ: gizzard, jejunum, caecum and colon (n=10/group). The groups were compared using nonparametric Kruskal-Wallis one-way analysis of variance followed by nonparametric Mann-Whitney test. Statistical analysis from experiment 2 was not performed due to the small sample size. Kruskal-Wallis one-way analysis and nonparametric Mann-Whitney *U* tests were performed using GraphPad Prism version 6 and *p*-values <0.05 were considered significant. [N.B. All statistical analysis for experiment 1 was performed by Clara Atterby, Uppsala University, Uppsala, Sweden.]

Ethics statement

All animal experiments were conducted in accordance with regulations provided by the Swedish Board of Agriculture and were approved by the Ethical Committee on Animal Experiments in Uppsala (permit number C20/14). Collection of *Campylobacter* isolates from wild birds at Ottenby Bird Observatory was approved by the Animal Experimentation Committee of Linköping (permit number Dnr112-11).

References

Axelsson-Olsson, D., Waldenström, J., Broman, T., Olsen, B. and Holmberg, M., 2005. Protozoan Acanthamoeba polyphaga as a Potential Reservoir for Campylobacter jejuni. *Applied and Environmental Microbiology*, 71(2), pp.987–992.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A. and Zagnitko, O., 2008. The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics*, 9(1), p.75.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. a. and Pevzner, P. a., 2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*, 19(5), pp.455–477.

Baumler, A. and Fang, F.C., 2013. Host Specificity of Bacterial Pathogens. *Cold Spring Harbor Perspectives in Medicine*, 3(12), pp.a010041–a010041.

Bereswill, S., Fischer, A., Plickert, R., Haag, L.-M., Otto, B., Kühl, A.A., Dashti, J.I., Zautner, A.E., Muñoz, M., Loddenkemper, C., Groß, U., Göbel, U.B. and Heimesaat, M.M., 2011. Novel Murine Infection Models Provide Deep Insights into the “Ménage à Trois” of Campylobacter jejuni, Microbiota and Host Innate Immunity. *PLoS ONE*, 6(6), p.e20953.

Brennhovd, O., Kapperud, G. and Langeland, G., 1992. Survey of thermotolerant Campylobacter spp. and Yersinia spp. in three surface water sources in Norway. *International Journal of Food Microbiology*, 15(3–4), pp.327–338.

Broman, T., Waldenström, J., Dahlgren, D., Carlsson, I., Eliasson, I. and Olsen, B., 2004. Diversities and similarities in PFGE profiles of Campylobacter jejuni isolated from migrating birds and humans. *Journal of applied microbiology*, 96(4), pp.834–43.

Chaloner, G., Wigley, P., Humphrey, S., Kemmett, K., Lacharme-Lora, L.,

Humphrey, T. and Williams, N., 2014. Dynamics of dual infection with Campylobacter jejuni strains in chickens reveals distinct strain-to-strain variation in infection ecology. *Applied and environmental microbiology*, 80(20), pp.6366–72.

Colles, F.M., Ali, J.S., Sheppard, S.K., McCarthy, N.D. and Maiden, M.C.J., 2011. Campylobacter populations in wild and domesticated Mallard ducks (Anas platyrhynchos). *Environmental microbiology reports*, 3(5), pp.574–580.

Colles, F. M., Dingle, K.E., Cody, A.J. and Maiden, M.C.J., 2008a. Comparison of Campylobacter populations in wild geese with those in starlings and free-range poultry on the same farm. *Applied and environmental microbiology*, 74(11), pp.3583–90.

Colles, Frances M., Jones, T.A., McCarthy, N.D., Sheppard, S.K., Cody, A.J., Dingle, K.E., Dawkins, M.S. and Maiden, M.C.J., 2008b. Campylobacter infection of broiler chickens in a free-range environment. *Environmental Microbiology*, 10(8), pp.2042–2050.

Connor, T.R., Loman, N.J., Thompson, S.E.-, Smith, A., Southgate, J., Poplawski, R., Bull, M.J., Richardson, E., Ismail, M., Thompson, S.E.-, Kitchen, C., Guest, M., Bakke, M., Sheppard, S.K., Pallen, M.J., Elwood-Thompson, S., Kitchen, C., Guest, M., Bakke, M., Sheppard, S.K. and Pallen, M.J., 2016. CLIMB (the Cloud Infrastructure for Microbial Bioinformatics): an online resource for the medical microbiology community. *Microbial Genomics*, 2(9), p.064451.

Dearlove, B.L., Cody, A.J., Pascoe, B., Méric, G., Wilson, D.J. and Sheppard, S.K., 2016. Rapid host switching in generalist Campylobacter strains erodes the signal for tracing human infections. *The ISME journal*, 10(3), pp.721–9.

Dicksved, J., Ellström, P., Engstrand, L. and Rautelin, H., 2014. Susceptibility to Campylobacter infection is associated with the species composition of the human fecal microbiota. *mBio*, 5(5), pp.e01212-14.

Dingle, K.E., Colles, F.M., Ure, R., Wagenaar, J.A., Duim, B., Bolton, F.J., Fox, A.J., Wareing, D.R.A. and Maiden, M.C.J., 2002. Molecular Characterization of Campylobacter jejuni Clones: A Basis for Epidemiologic Investigation. *Emerging Infectious Diseases*, 8(9), pp.949–955.

Dingle, K.E., Colles, F.M., Wareing, D.R.A., Ure, R., Fox, A.J., Bolton, F.E., Bootsma, H.J., Willems, R.J.L., Urwin, R. and Maiden, M.C.J., 2001. Multilocus Sequence Typing System for Campylobacter jejuni. *Journal of Clinical Microbiology*, 39(1), pp.14–23.

Food, E., Authority, S. and Centre, E., 2014. The European Union Summary Report on Trends and Sources of Zoonoses , Trends and Sources of Zoonoses , Zoonotic Agents and Food-borne Outbreaks in 2012. 12(2).

Glunder, G., 1995. [Infectivity of Campylobacter jejuni and Campylobacter coli in chickens]. *Berl Munch Tierarztl Wochenschr* [Online], 108(3), pp.101–4.

Griekspoor, P., Colles, F.M., McCarthy, N.D., Hansbro, P.M., Ashhurst-Smith, C., Olsen, B., Hasselquist, D., Maiden, M.C.J. and Waldenström, J., 2013. Marked host specificity and lack of phylogeographic population structure of Campylobacter jejuni in wild birds. *Molecular ecology*, 22(5), pp.1463–72.

Griekspoor, P., Engvall, E.O., Åkerlind, B., Olsen, B. and Waldenström, J., 2015. Genetic diversity and host associations in Campylobacter jejuni from human cases and broilers in 2000 and 2008. *Veterinary Microbiology*, 178(1–2), pp.94–98.

Gurevich, A., Saveliev, V., Vyahhi, N. and Tesler, G., 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, 29(8), pp.1072–1075.

Haag, L.-M., Fischer, A., Otto, B., Plickert, R., Kühl, A.A., Göbel, U.B., Bereswill, S. and Heimesaat, M.M., 2012. Intestinal Microbiota Shifts towards Elevated Commensal Escherichia coli Loads Abrogate Colonization Resistance against Campylobacter jejuni in Mice. *PLoS ONE*, 7(5), p.e35988.

de Haan, C.P.A., Kivisto, R., Hakkinen, M., Rautelin, H. and Hanninen, M.L., 2010. Decreasing Trend of Overlapping Multilocus Sequence Types between Human and Chicken Campylobacter jejuni Isolates over a Decade in Finland. *Applied and Environmental Microbiology*, 76(15), pp.5228–5236.

Hänel, I., Müller, J., Müller, W. and Schulze, F., 2004. Correlation between invasion of Caco-2 eukaryotic cells and colonization ability in the chick gut in Campylobacter jejuni. *Veterinary Microbiology*, 101(2), pp.75–82.

Humphrey, T., O'Brien, S. and Madsen, M., 2007. Campylobacters as zoonotic

pathogens: A food production perspective. *International Journal of Food Microbiology*, 117(3), pp.237–257.

Järhult, J.D., Muradrasoli, S., Wahlgren, J., Söderström, H., Orozovic, G., Gunnarsson, G., Bröjer, C., Latorre-Margalef, N., Fick, J., Grabic, R., Lennerstrand, J., Waldenström, J., Lundkvist, A. and Olsen, B., 2011. Environmental levels of the antiviral oseltamivir induce development of resistance mutation H274Y in influenza A/H1N1 virus in mallards. *PloS one*, 6(9), p.e24742.

Jolley, K.A., Bray, J.E. and Maiden, M.C.J., 2018. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Research*, 3(0), p.124.

Kapperud, G. and Rosef, O., 1983. Avian wildlife reservoir of Campylobacter fetus subsp. jejuni, Yersinia spp., and Salmonella spp. in Norway. *Applied and environmental microbiology*, 45(2), pp.375–80.

Katoh, K., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30(14), pp.3059–3066.

Koressaar, T. and Remm, M., 2007. Enhancements and modifications of primer design program Primer3. *Bioinformatics*, 23(10), pp.1289–1291.

Korolik, V., Alderton, M.R., Smith, S.C., Chang, J. and Coloe, P.J., 1998. Isolation and molecular analysis of colonising and non-colonising strains of Campylobacter jejuni and Campylobacter coli following experimental infection of young chickens. *Veterinary microbiology*, 60(2–4), pp.239–49.

McCrea, B.A., Tonooka, K.H., VanWorth, C., Atwill, E.R. and Schrader, J.S., 2006. Colonizing capability of Campylobacter jejuni genotypes from low-prevalence avian species in broiler chickens. *Journal of food protection*, 69(2), pp.417–20.

Méric, G., Miragaia, M., de Been, M., Yahara, K., Pascoe, B., Mageiros, L., Mikhail, J., Harris, L.G., Wilkinson, T.S., Rolo, J., Lamble, S., Bray, J.E., Jolley, K.A., Hanage, W.P., Bowden, R., Maiden, M.C.J.J., Mack, D., de Lencastre, H., Feil, E.J., Corander, J., Sheppard, S.K., Meric, G., Miragaia, M., de Been, M., Yahara, K., Pascoe, B., Mageiros, L., Mikhail, J., Harris,

L.G., Wilkinson, T.S., Rolo, J., Lamble, S., Bray, J.E., Jolley, K.A., Hanage, W.P., Bowden, R., Maiden, M.C.J.J., Mack, D., de Lencastre, H., Feil, E.J., Corander, J. and Sheppard, S.K., 2015. Ecological Overlap and Horizontal Gene Transfer in Staphylococcus aureus and Staphylococcus epidermidis. *Genome Biology and Evolution*, 7(5), pp.1313–1328.

Méric, G., Yahara, K., Mageiros, L., Pascoe, B., Maiden, M.C.J., Jolley, K.A. and Sheppard, S.K., 2014. A reference pan-genome approach to comparative bacterial genomics: identification of novel epidemiological markers in pathogenic Campylobacter. *PloS one*, 9(3), p.e92798.

Mohan, V., Stevenson, M., Marshall, J., Fearnhead, P., Holland, B.R., Hotter, G. and French, N.P., 2013. Campylobacter jejuni colonization and population structure in urban populations of ducks and starlings in New Zealand. *MicrobiologyOpen*, 2(4), pp.659–73.

Parkhill, J., Wren, B.W., Mungall, K., Ketley, J.M., Churcher, C., Basham, D., Chillingworth, T., Davies, R.M., Feltwell, T., Holroyd, S., Jagels, K., Karlyshev, A. V., Moule, S., Pallen, M.J., Penn, C.W., Quail, M.A., Rajandream, M.A., Rutherford, K.M., van Vliet, A.H., Whitehead, S. and Barrell, B.G., 2000. The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences. *Nature*, 403(6770), pp.665–8.

Pascoe, B., Méric, G., Murray, S., Yahara, K., Mageiros, L., Bowen, R., Jones, N.H., Jeeves, R.E., Lappin-Scott, H.M., Asakura, H. and Sheppard, S.K., 2015. Enhanced biofilm formation and multi-host transmission evolve from divergent genetic backgrounds in C ampylobacter jejuni. *Environmental Microbiology*, 17(11), pp.4779–4789.

Petersen, L., Nielsen, E.M., Engberg, J., On, S.L.W. and Dietz, H.H., 2001. Comparison of Genotypes and Serotypes of Campylobacter jejuni Isolated from Danish Wild Mammals and Birds and from Broiler Flocks and Humans. *Applied and Environmental Microbiology*, 67(7), pp.3115–3121.

Roumagnac, P., Weill, F.-X., Dolecek, C., Baker, S., Brisse, S., Chinh, N.T., Le, T.A.H., Acosta, C.J., Farrar, J., Dougan, G. and Achtman, M., 2006. Evolutionary history of Salmonella typhi. *Science (New York, N.Y.)*, 314(5803), pp.1301–4.

Selander, R.K., Beltran, P., Smith, N.H., Helmuth, R., Rubin, F.A., Kopecko, D.J.,

Ferris, K., Tall, B.D., Cravioto, A. and Musser, J.M., 1990. Evolutionary genetic relationships of clones of Salmonella serovars that cause human typhoid and other enteric fevers. *Infection and immunity*, 58(7), pp.2262–75.

Sheppard, S.K., Cheng, L., Méric, G., de Haan, C.P.A., Llarena, A.-K., Marttinen, P., Vidal, A., Ridley, A., Clifton-Hadley, F., Connor, T.R., Strachan, N.J.C., Forbes, K., Colles, F.M., Jolley, K.A., Bentley, S.D., Maiden, M.C.J., Hänninen, M.-L., Parkhill, J., Hanage, W.P. and Corander, J., 2014. Cryptic ecology among host generalist *Campylobacter jejuni* in domestic animals. *Molecular Ecology*, 23(10), pp.2442–2451.

Sheppard, S.K., Colles, F.M., McCarthy, N.D., Strachan, N.J.C., Ogden, I.D., Forbes, K.J., Dallas, J.F. and Maiden, M.C.J., 2011. Niche segregation and genetic structure of Campylobacter jejuni populations from wild and agricultural host species. *Molecular ecology*, 20(16), pp.3484–90.

Sheppard, S.K., Dallas, J.F., MacRae, M., McCarthy, N.D., Sproston, E.L., Gormley, F.J., Strachan, N.J.C., Ogden, I.D., Maiden, M.C.J. and Forbes, K.J., 2009a. Campylobacter genotypes from food animals, environmental sources and clinical disease in Scotland 2005/6. *International Journal of Food Microbiology*, 134(1–2), pp.96–103.

Sheppard, S.K., Dallas, J.F., Strachan, N.J.C., MacRae, M., McCarthy, N.D., Wilson, D.J., Gormley, F.J., Falush, D., Ogden, I.D., Maiden, M.C.J. and Forbes, K.J., 2009b. Campylobacter Genotyping to Determine the Source of Human Infection. *Clinical Infectious Diseases*, 48(8), pp.1072–1078.

Sheppard, S.K., Didelot, X., Meric, G., Torralbo, A., Jolley, K. a, Kelly, D.J., Bentley, S.D., Maiden, M.C.J., Parkhill, J. and Falush, D., 2013. Genome-wide association study identifies vitamin B5 biosynthesis as a host specificity factor in Campylobacter. *Proceedings of the National Academy of Sciences*, 110(29), pp.11923–11927.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 22(21), pp.2688–2690.

Stas, T., Jordan, F.T. and Woldehiwet, Z., 1999. Experimental infection of chickens with Campylobacter jejuni: strains differ in their capacity to colonize the intestine. *Avian pathology : journal of the W.V.P.A*, 28(1), pp.61–4.

Taylor, L.H., Latham, S.M. and Woolhouse, M.E.J., 2001. Risk factors for human disease emergence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 356(1411), pp.983–989.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. and Rozen, S.G., 2012. Primer3-new capabilities and interfaces. *Nucleic Acids Research*, 40(15), pp.1–12.

Waldenström, J., Axelsson-Olsson, D., Olsen, B., Hasselquist, D., Griekspoor, P., Jansson, L., Teneberg, S., Svensson, L. and Ellström, P., 2010. Campylobacter jejuni colonization in wild birds: results from an infection experiment. *PloS one*, 5(2), p.e9082.

Waldenström, J., Broman, T., Carlsson, I., Hasselquist, D., Achterberg, R.P., Wagenaar, J.A. and Olsen, B., 2002. Prevalence of Campylobacter jejuni, Campylobacter lari, and Campylobacter coli in different ecological guilds and taxa of migrating birds. *Applied and environmental microbiology*, 68(12), pp.5911–7.

Waldenström, J. and Griekspoor, P., 2014. Ecology and Host Association in Campylobacter in Wild Birds. *Campylobacter Ecology and Evolution*, [Online]. Caister Academic Press: Norfolk, pp.265–284.

Whiley, H., van den Akker, B., Giglio, S. and Bentham, R., 2013. The Role of Environmental Reservoirs in Human Campylobacteriosis. *International Journal of Environmental Research and Public Health*, 10(11), pp.5886–5907.

Wickham, H., 2009. ggplot2. *Springer-Verlag New York* [Online].

Woolhouse, M.E.J., Taylor, L.H. and Haydon, D.T., 2001. Population Biology of Multihost Pathogens. *Science*, 292(5519), pp.1109–1112.

Yahara, K., Méric, G., Taylor, A.J., de Vries, S.P.W., Murray, S., Pascoe, B., Mageiros, L., Torralbo, A., Vidal, A., Ridley, A., Komukai, S., Wimalarathna, H., Cody, A.J., Colles, F.M., McCarthy, N., Harris, D., Bray, J.E., Jolley, K.A., Maiden, M.C.J., Bentley, S.D., Parkhill, J., Bayliss, C.D., Grant, A., Maskell, D., Didelot, X., Kelly, D.J. and Sheppard, S.K., 2017. Genome-wide association of functional traits linked with Campylobacter jejuni survival from farm to fork. *Environmental microbiology*, 19(1), pp.361–380.

Zahrt, T.C., 1998. *A Genetic Analysis of the Salmonella Typhi Host Range*.

University of Illinois, Urbana-Champaign.

Data availability

Genomes sequenced as part of this chapter are archived on the NCBI associated BioProject accession PRJNA415188. Contiguous assemblies of all genome sequences compared are available at the public data repository Figshare (doi.org/10.6084/m9.figshare.12739991).

Chapter 6

Discussion

Discussion

The work presented in this dissertation has aimed to advance knowledge of the genomics and evolution of bacteria of the genus *Campylobacter*. Although each chapter has its' own hypothesis and narrative, all chapters investigate the specific adaptation of bacterial species, lineages or genes in different hosts and environments. Chapter 2 focuses on different species of the *Campylobacter* genus describing their population structure. It further investigates the extent of genetic exchange between isolates of different species by studying the variation on the core and accessory genome and by examining the effect of recombination. Chapter 3 is concerned with the emergence of a cattle-associated lineage within the most well-studied *Campylobacter* species (*C. jejuni*). As in chapter 2, chapter 3 investigates this emergence by studying genome variation (core and accessory) and the effect of recombination. Chapter 4 is focused on the circulation of AMR genes among *Campylobacter* bacteria isolated from humans, animals and the environment. Finally, chapter 5 addresses the use of genomics as a tool for identifying lineage-specific genes as molecular markers to monitor bacterial load in *in vivo* animal experiments. A common theme of this dissertation is to answer questions related to host adaptation and bacterial evolution by studying the genetic variation on different levels (species, lineages, genes). In this chapter, I discuss the most compelling results from each chapter by placing them in a broader context to enhance our understanding of the evolution of those versatile pathogens.

The effect of anthropogenic change on the evolution of *Campylobacter*

Pathogen spill-over into a new host population is associated with emergence of various zoonotic pathogens, including bacteria, and has been linked with anthropogenic environmental changes on the planet, including deforestation, destruction of wildlife habitats and intensive livestock production (McMichael, 2004; Woolhouse et al., 2005). The dramatic rise in numbers of agricultural animals has expanded the host niches of *Campylobacter*, favouring the emergence and expansion of agriculture-associated species and lineages (Thépault et al., 2017; Sheppard et al., 2013; Morley et al., 2015). In chapter 2, exploring the isolation source of hundreds of *Campylobacter* bacteria, showed that around 25% of the genus species are associated with livestock animals (chickens, cattle, pigs). Other studies have highlighted the increased opportunities for transfer of bacterial pathogens between humans and animals driven by the recent intensification

livestock farming (Richardson et al., 2018; Weinert et al., 2015). Similarly, in chapter 3, the dramatic rise in numbers of cattle globally has coincided with the emergence of cattle-associated lineages. Dating analysis has placed the emergence of cattle specialists *C. jejuni* between 1850 and 20th century. The emergence of new *Campylobacter* species and lineages in farmed animals represents a risk for infecting human populations through the food production chain. The availability of WGS of those isolates grants the opportunity to investigate the genomics and evolution of *Campylobacter*.

Microbial infections have become an increased threat to public health due to the dramatic rise of AMR in recent years (Balouiri et al., 2016). The use of antibiotics in human and veterinary medicine as well as in agricultural animals as growth promoters has been linked with increasing AMR trends in various pathogenic bacteria (Schwarz et al., 2001; Teuber, 2001; Engberg et al., 2001; Livermore, 2007). *Campylobacter* is one of the priority pathogens according to the WHO with an increasing trend towards the presence of MDR in *C. coli* isolates (WHO, 2017). The work presented in chapter 4 provides evidence of high AMR levels against ciprofloxacin, the drug of choice against campylobacteriosis infections, along with tetracycline, in *C. jejuni* and *C. coli* strains isolated from various sources in Spain. Furthermore, consistent with previous studies (Luangtongkum et al., 2009; Pascoe et al., 2017; Food and Authority, 2019), analysis in chapter 4, shows an increase of MDR in *C. coli* isolates. More worryingly, resistance against erythromycin, the alternative drug in treating clinical campylobacteriosis, was detected in *C. coli*. Work presented in chapter 2, showed that many AMR genes were present in *Campylobacter* species isolated from agricultural animals. These results show that AMR is widespread among *Campylobacter* isolates from livestock animals.

The impact of HGT on host adaptation

Many theories have tried to explain the genetic variation observed in bacteria in the microbial world. Scientists have been debating on whether bacteria could be classified as species and under which concepts and definitions (Doolittle and Zhaxybayeva, 2009). Ongoing sampling from various sources and the development of new sophisticated software and pipelines in genomics has aided to in-depth exploration of the genetic diversity of various bacterial microorganisms and addressed complicating questions related to HGT. Bacteria that belong to the same

species are known to exchange DNA via HGT, but the extent of gene flow between related species when sharing or not the same host is yet to be determined.

HGT has a fundamental role in the evolution of recombinogenic bacteria and, specifically *Campylobacter* where it generates diversity at twice the rate of *de novo* mutation (Wilson et al., 2009). However, the extent to which recombination mediates gene transfer within and between *Campylobacter* species is not well understood (Wilson et al., 2009). Work in chapter 2, takes advantage of the availability and broad host range of hundreds of *Campylobacter* genomes to investigate the extent of genetic exchange between species that cohabit the same host or environment. Recombination analysis in chapter 2, quantified the DNA that was transferred between strains from different species that may or may not found in the same host. The analysis revealed specific donor-recipient species pairs with enhanced recombination within rather than between a certain niche. The work in chapter 2 represents a novel approach to quantifying the proportion of the genome that is recombining. Within-host recombination was ~0.6 times greater than between host in some species pairs. Mapping the DNA segments involved in recombination revealed hot spots of recombination in the genome. The putative functions of the genes involved in those hot spots add useful knowledge on potentially essential mechanisms for adaptation in niches and environments.

In chapter 3, a detailed recombination analysis in *C. jejuni* was carried out to investigate the genomic changes on the emergence of cattle specialist lineages. The effect of recombination in generating diversity in cattle specialists (ST-61) was high. Analysis in chapter 3 using ClonalFrameML inferred an *r*/*m* ratio much higher than that of the generalist ST-21 or the entire dataset. Lower estimates for *r*/*m* have been reported for *C. jejuni* (Vos and Didelot, 2009) than what was inferred in chapter 3, but this is probably because these estimates were based on seven MLST genes (Wilson et al., 2009). Recombination might have a weak effect on the highly conserved MLST genes compared to the whole genome. The recombination analysis revealed segments of DNA that were recombining and characterized their function. Genes were related to metabolic functions, cell envelope biogenesis and cell motility.

Gene loss and genotype-phenotype interactions in *C. jejuni*

Previous studies have shown that bacteria occupying specific niches can display reductive evolution as a result of genetic drift or linkage to beneficial mutations that are not necessarily adaptive (Batut et al., 2014, Sheppard et al., 2018). However, gene loss in host specialist lineages has been observed in *Campylobacter* (Morley et al., 2015) and *S. enterica* (Koskiniemi et al., 2012). Consistent with these studies, work in chapter 3 revealed significant gene loss associated with the emergence of *C. jejuni* cattle specialists. The gene loss included a genomic island encoding the flagellin O-linked glycosylation system. Homoplasy analysis showed that this gene block was also absent in other divergent *C. jejuni* cattle specialists indicating a beneficial advantage of those strains in adaptation in the cattle niche.

One of the challenges in biology is to mechanistically link genotypes with phenotypes. Comparative genomics analysis in chapter 3, revealed homoplasious genomic signatures associated with adaptive evolution and linked their function with phenotype assays in a collection of *C. jejuni* bacterial strains. *In vitro* phenotype assays showed that glycosylation gene block loss had phenotypic changes associated with reductions in cell hydrophobicity, autoagglutination and biofilm formation in wild-type cattle compared to chicken specialist strains. Further *in vivo* experiments in cattle would shed more light on how gene loss is influencing adaptation of cattle-specialist *C. jejuni* bacteria in cattle.

AMR is highly distributed among *Campylobacter* species associated with agriculture

HGT is linked with the widespread distribution of AMR genes across bacterial species (de la Cruz and Davies, 2000). Work presented in chapters 2, 3 and 4 focused on the effect of HGT on large numbers of genomes. Antibiotics represent a very strong selective pressure to which bacteria need to rapidly adapt. The spread of AMR genes among human, animals and the environment via HGT on a global scale has been discussed before (Mazel and Davies, 1999). However, there is limited information about the transmission dynamics of AMR among animals, humans and the environment in the *Campylobacter* genus. Pangenomic analysis in chapter 2, showed the presence of multiple AMR genes in six agricultural-associated *Campylobacter* species. Additionally, some AMR genes were found in species associated with marine mammals, wild birds and environmental waters.

Studies have reported AMR gene clusters which can be transferred between lineages and species via HGT (Qin et al., 2012; Derbise et al., 1996; Derbise et al., 1997; Baker et al., 2018). The work presented in chapter 4, identified localization of AMR gene clusters on GIs. Several of those GIs were located on plasmids or integrative conjugative elements indicating spread of gene clusters via HGT. Syntenic arrangements of AMR genes conferring resistance to multiple classes of antibiotics have been described before (Werner et al., 2003; Derbise et al., 1996; Derbise et al., 1997). Work in chapter 4 shows similar gene localizations as well as the presence of a new gene association among isolates from different sources and *Campylobacter* species. Similar results were obtained in chapter 2, where gene associations were identified in different *Campylobacter* species, indicating the dissemination of AMR genes via HGT. These results suggest that strong selection has enabled the dissemination of AMR genes in different *Campylobacter* species inhabiting livestock animals with a further spill-over to humans and the environment.

Host-specific gene pools

Bacterial species that are found in multiple hosts carry genetic signatures of host adaptation in their genomes (Toft and Andersson, 2010). These signatures are the result of diversification from the ancestral gene pool and/or genetic elements that confer an advantage for bacteria to adapt and survive in the host niche. Identifying those signatures in the genome can reveal gene functions and important biological mechanisms underlying adaptation associated with a jump into a new host (Sheppard et al., 2018). However, it is challenging to differentiate which genetic signatures are the result of adaptation and which the result of genetic bottleneck and drift.

The work presented in chapter 2 provides a detailed within-host recombination analysis for *Campylobacter* species that share the same host. The analysis identified SNPs with a high probability of recombining between a donor and a recipient strain and mapped highly recombinant genes involved in different comparisons. Analysis of different *Campylobacter* species inhabiting the chicken niche, identified an AMR gene, *gyrA*, as highly recombinant. *C. jejuni* and *C. coli* strains isolated from chickens are highly resistant against ciprofloxacin (Food and Authority, 2019) with mutation in the *gyrA* gene conferring resistance against that antibiotic (Hormeño et al., 2016). Further work in chapter 2 detected high levels of

recombination in *Campylobacter* species inhabiting the cattle niche. Those species included strains of a cattle-specialist *C. jejuni* lineage (ST-61). This is consistent with work in chapter 3, where this lineage was found to be highly recombinogenic. Owing to their unique digestive system, cattle represent a distinct habitat for large populations of microbes (Flemming and Wuertz, 2019), which may act as a microcosm of interactions between *Campylobacter* species. The work presented in chapter 2 identified a gene (*glmS*) with a transposon specific site located just downstream of this gene. Transposons are mobile genetic elements known to transfer between different bacterial strains carrying various genes (Choi, 2009). Finally, recombination analysis focusing on species that are found in pigs and cattle detected gene *napA* to be very recombinogenic. This gene encodes a key nitrate reductase enzyme that allows *Campylobacter* to grow under microaerobic conditions (Pittman et al., 2007) and may help balance the rate of biosynthesis with the varying supply of potentially toxic (to cattle and pigs) nitrite (Alexander et al., 2009).

The work presented in chapter 3, identified genes and alleles that are associated with adaptive changes by focusing on homoplasious changes that had occurred in other divergent cattle specialists. Similar approach was performed in another study on *S. aureus*, to differentiate adaptive changes from the ones that occur as a result of a bottleneck or genetic drift from the ancestral gene pool (Murray et al., 2017). Homoplasy analysis revealed genes involved in diverse functions including adhesion to epithelial cells and biosynthesis of thiamine and molybdenum. A recent study identified a gene encoding for a fibronectin-binding protein which is associated with host colonization (Smith et al., 2020). Work in chapter 3, identified a putative fibronectin/fibrinogen-binding protein gene as homoplasious in cattle specialist lineages likely involved in host cell adhesion in cattle. Other homoplasious genetic elements included genes involved in thiamine and molybdenum biosynthesis. Metabolic functions of those genes might be related with differences in host diet between chickens and cattle. Although speculations about the putative functions of these genes link them with differences in host anatomy and physiology, further *in vitro* studies are needed to confirm their adaptive role.

Lineage-specific accessory genes and their use as molecular markers

Monitoring the colonization ability of different bacterial strains in *in vivo* animal experiments can be performed by culturing the bacteria directly from the sample on selective media and counting the CFUs. While this is true for single infection experiments, it is impossible to differentiate multiple bacterial strains of the same species in competition experiments. The work presented in chapter 5 shows how genomics can be used to identify unique genes that can be used as targets for downstream analysis such as real-time PCR, to measure strain specific colonization ability. Single and competition infection experiments in mallards were performed with *C. jejuni* strains isolated from three different bird species. All three strains were whole genome sequenced and belonged to three clonal complexes. Two clonal complexes were wild bird specialists, one of ducks, and one of song thrushes, while the third a generalist clonal complex from chickens. Pangenomic analysis of those three strains used in infection experiments augmented with publicly available isolates from the same ST-complexes revealed accessory genes unique to each ST-complex. Those genes appear to be ST-complex specific and, due to the strong host link between *C. jejuni* ST-complexes with a particular host, potentially host specific. It is possible that some of those genes might be the reason underlying differences in colonization ability of strains used in the infection experiments. However, a larger number of strains is required for *in vitro* phenotype assays and *in vivo* infection experiments to draw more generalizable conclusions.

The mobilome of the *Campylobacter* genus

Many bacterial species are found in the same environments. These bacteria have different genetic backgrounds but require the same metabolic functions to adapt and survive. Recombination allows for rapid adaptation to a host environment facilitating the spread of adaptive genes through bacteria populations. Consistent with this, work presented in chapter 4, showed the same AMR alleles shared between isolates sampled from multiple sources and different *Campylobacter* species. Recombination analysis in chapter 2 identified genes that were recombining in multiple *Campylobacter* species. These genes were associated with MDR highlighting the potential for selective pressure imposed by antibiotics in multiple host niches. Other genes with evidence for interspecies HGT included those involved in biosynthesis of a polysaccharide and a hydrogenase that utilizes nickel, a metal that is present in the environment. Further functional characterization and

in vitro phenotype assays are required to define the mechanistic functions of these genes and explain why they are recombining in multiple *Campylobacter* species.

Conclusions and future work

The work presented in this thesis has shown the impact of intensive livestock production on the evolution of agricultural-associated *Campylobacter* species, particularly on *C. jejuni*, where the expanded niche of cattle in the last century, favoured the emergence and dissemination of cattle specialist lineages. Work in this thesis shows that HGT plays a key role in facilitating intra- and interspecies genetic variation in the *Campylobacter* genus. Findings revealed various genes that potentially promote host-adaptive evolution. The work in this thesis improves understanding of the genetic and functional basis of host adaptation in *Campylobacter*. Future work will focus on testing the mechanistic basis of genes identified from genomic analysis using *in vitro* phenotype assays. Further work will include *in vivo* animal experiments to confirm the adaptive role of those genes.

Appendix

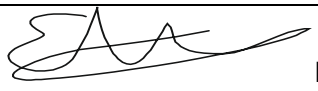
Appendix for chapter 2

2.1. Supplementary form SF1

2.2. Supplementary figures

2.3. Supplementary tables

2.1. Supplementary form SF1.

This declaration concerns the article entitled:			
Interspecies recombination in agricultural *Campylobacter* is influenced by the song (gene) and the singer (strain)			
Publication status (tick one) **Draft manuscript** [ ] **Submitted** [X] **In review** [ ] **Accepted** [ ] **Published** [ ]			
Publication details (reference)			
Copyright status (tick the appropriate statement) I hold the copyright for this material [X] Copyright is retained by the publisher, but I have been given permission to replicate the material here [ ]			
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	The candidate contributed to / considerably contributed to / predominantly executed the… Formulation of ideas: Predominantly executed the 90% Design of methodology: Predominantly executed the 90% Experimental work: Predominantly executed the 95% Presentation of data in journal format: Predominantly executed the 90%		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed	 Evangelos Mourkas	**Date**	31/07/2020

2.2. Supplementary figures

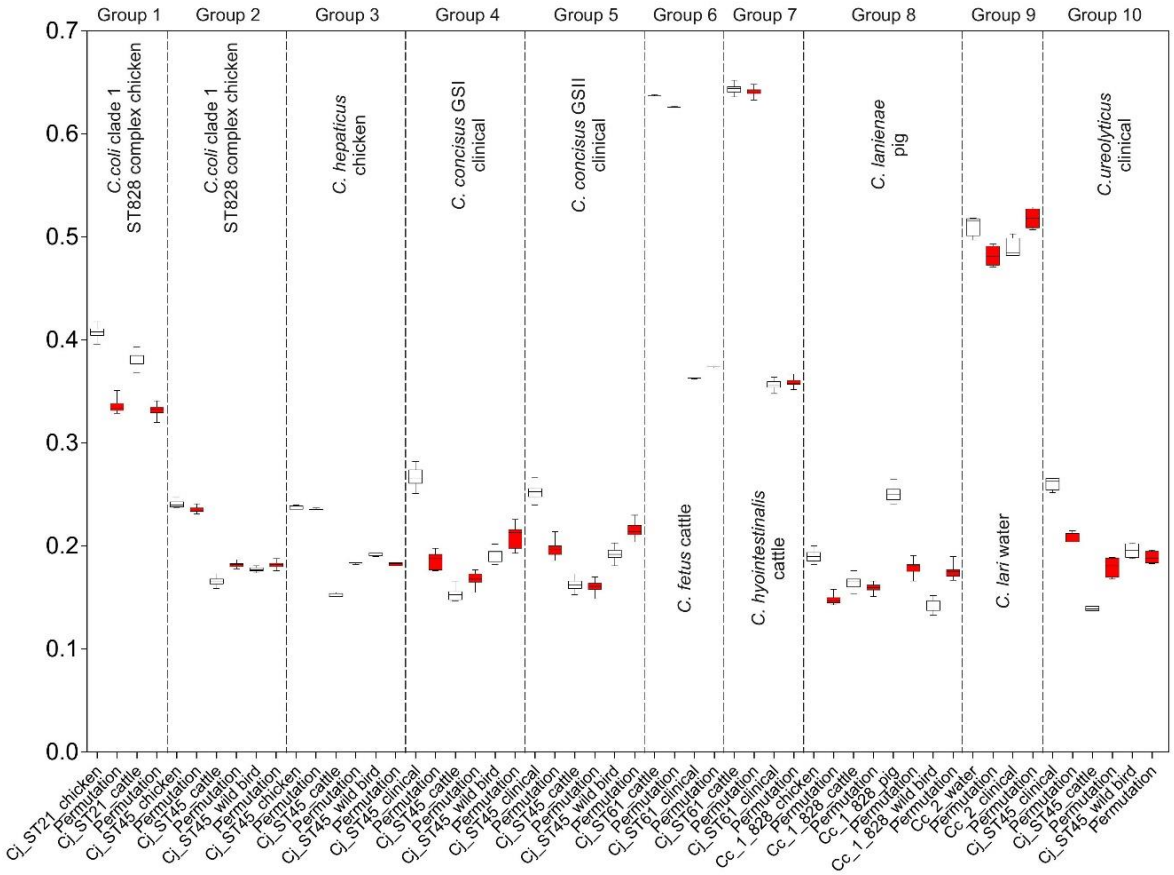


Figure S2.1. Probability of the recipient genomes sharing DNA with each donor groups is illustrated as box whiskers (white) for every donor-recipient comparison for all ten groups that supported our hypothesis. The analysis where the host data were randomized across all isolates is illustrated as box whiskers (red). The probability of copying DNA from a donor to a recipient genome is shown on the *y* axis. The midline in the box whiskers indicates the mean and the error bars the standard deviation.

2.3. Supplementary tables

All supplementary tables are available at the public data repository FigShare (doi.org/10.6084/m9.figshare.12741296).

Table S2.1. Details of isolates used in chapter 2.

Table S2.2. Within-host highly (>95th percentile) recombining genes.

Table S2.3. Recombination parameters as calculated by ClonalFrameML.

Table S2.4. Genes involved in interspecies recombination in 10 group comparisons.

Appendix for chapter 3

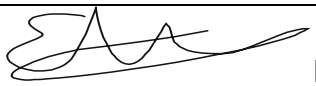
3.1. Supplementary form SF2

3.2. Supplementary figures

3.3. Supplementary tables

3.4. Supplementary references

3.1. Supplementary form SF2.

This declaration concerns the article entitled:			
Agricultural intensification and the evolution of host specialism in the enteric pathogen *Campylobacter jejuni*			
Publication status (tick one) **Draft manuscript** [ ] **Submitted** [ ] **In review** [ ] **Accepted** [ ] **Published** [X]			
Publication details (reference)	Mourkas, E., Taylor, J.T., Méric, G., Bayliss, C.S., Pascoe, B., Mageiros, L., Calland, K.C., Hitchings, M.D., Ridley, A., Vidal, A., Forbes, J.K., Strachan, J.C.N., Parker, T.C., Parkhill, J., Jolley, A.K., Cody, J.A., Maiden, C.J.M., Kelly, J.D., Sheppard, S.K. 2020. Agricultural intensification and the evolution of host specialism in the enteric pathogen *Campylobacter jejuni*. *PNAS* 117 (20) 11018-11028. doi: 10.1073/pnas.1917168117.		
Copyright status (tick the appropriate statement) I hold the copyright for this material [X] Copyright is retained by the publisher, but I have been given permission to replicate the material here [ ]			
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	The candidate contributed to / considerably contributed to / predominantly executed the… Formulation of ideas: Considerably contributed to 70% Design of methodology: Predominantly executed the 95% Experimental work: Predominantly executed the 95% Presentation of data in journal format: Predominantly executed the 90%		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed	 Evangelos Mourkas	**Date**	31/07/2020

3.2. Supplementary figures

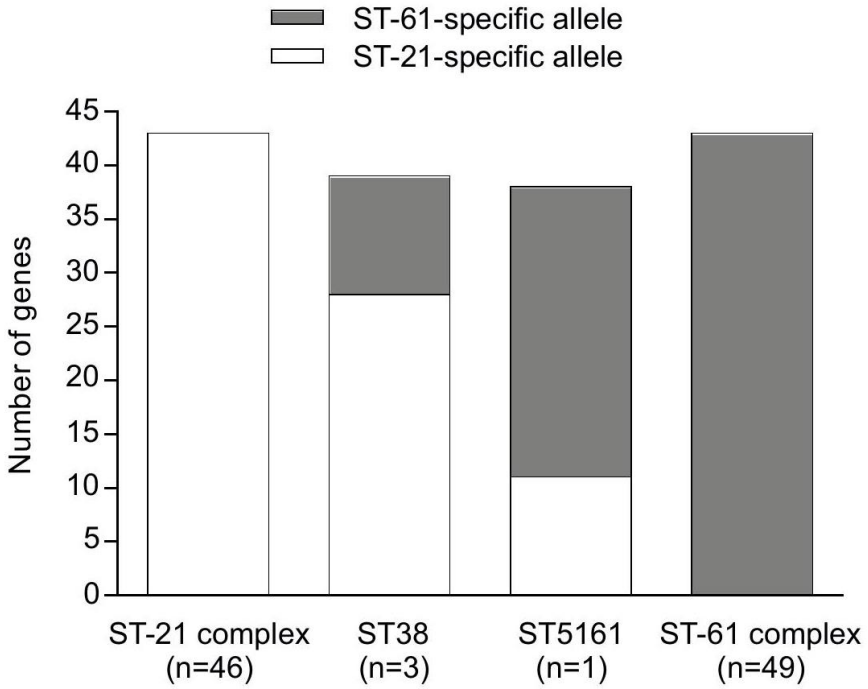


Figure S3.1. Distribution of genes that harbour ST-21 (white) or ST-61 (black) specific alleles. The number of genes is illustrated in the *y* axis while the ST-21 and ST-61 complexes as well as the intermediates ST38 and ST5161 in the *x* axis.

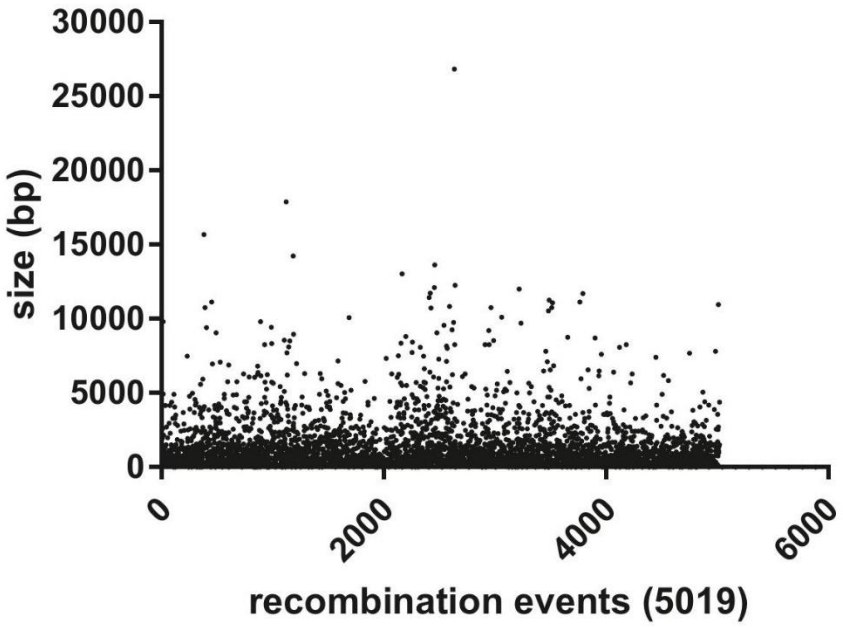


Figure S3.2. All recombination events as inferred by ClonalFrameML. The number of recombination events is plotted against the length (bp) of the recombinant sequence.

3.3. Supplementary tables

All supplementary tables are available at the public data repository FigShare (doi.org/10.6084/m9.figshare.12743339).

Table S3.1. Isolate information for the strains used in chapter 3.

Table S3.2. Prevalence in 13 *C. jejuni* clonal complexes of 301 genes identified as being distinctly associated with the emergence of ST-61 complex isolates from ST-21 clonal complex.

Table S3.3. Genetic elements associated with ST-42 complex cattle specialization.

Table S3.4. Primers and their corresponding sequences used for cloning and mutagenesis confirmation. The uppercase sequences of the geneblock primers are the adaptor regions used in the Gibson assembly cloning, while the lowercase sequences are the regions annealing to a region upstream of *cj1324* (F1) and just inside the *cj1324* coding region (R1) or towards the end of the *cj1332* coding region (F2) and downstream of *cj1332* (R2). The *KanF* and *KanR* primers are adaptors that also amplify the *kan* gene from pJMK30 (120).

Primer name	Sequence 5' – 3'
geneblockF1	GAGCTCGGTACCCGGGGATCCTCTAGAGTCgataaataccgcagaatgaat
geneblockR1	AAGCTGTCAAACATGAGAACCAAGGAGAATcacgacttttagcaaaaatataa
geneblockF2	GAATTGTTTTAGTACCTAGCCAAGGTGTGCtttgtagcaggagcttgtg
geneblockR2	AGAATACTCAAGCTTGCATGCCTGCAGGTCatgttttgattaaaaagctcttt
KanF	ATTCTCCTTGGTTCTCATGTTTGACAGCTTAT
KanR	GCACACCTTGGCTAGGTACTAAAACAATTCAT

3.4. Supplementary references

1. Atterby C, et al. (2018) The Potential of Isolation Source to Predict Colonization in Avian Hosts: A Case Study in *Campylobacter jejuni* Strains From Three Bird Species. *Front Microbiol* 9(March):1–12.

2. Baily JL, et al. (2015) Evidence of land-sea transfer of the zoonotic pathogen *Campylobacter* to a wildlife marine sentinel species. *Mol Ecol* 24(1):208–221.

3. Biggs PJ, et al. (2011) Whole-Genome Comparison of Two *Campylobacter jejuni* Isolates of the Same Sequence Type Reveals Multiple Loci of Different Ancestral Lineage. *PLoS One* 6(11):e27121.

4. Loc Carrillo C, et al. (2005) Bacteriophage Therapy To Reduce *Campylobacter jejuni* Colonization of Broiler Chickens. *Appl Environ Microbiol* 71(11):6554–6563.

5. Cody AJ, et al. (2013) Real-Time Genomic Epidemiological Evaluation of Human *Campylobacter* Isolates by Use of Whole-Genome Multilocus Sequence Typing. *J Clin Microbiol* 51(8):2526–2534.

6. Cooper KK, Cooper MA, Zuccolo A, Law B, Joens LA (2011) Complete Genome Sequence of *Campylobacter jejuni* Strain S3. *J Bacteriol* 193(6):1491–1492.

7. Gripp E, et al. (2011) Closely related *Campylobacter jejuni* strains from different sources reveal a generalist rather than a specialist lifestyle. *BMC Genomics* 12(1):584.

8. Gunther NW, Bono JL, Needleman DS (2015) Complete Genome Sequence of *Campylobacter jejuni* RM1285, a Rod-Shaped Morphological Variant. *Genome Announc* 3(6):e01361-15.

9. He Y, et al. (2015) Complete Genome Sequence of *Campylobacter jejuni* YH001 from Beef Liver, Which Contains a Novel Plasmid. *Genome Announc* 3(1):7–8.

10. Morley L, et al. (2015) Gene Loss and Lineage-Specific Restriction-Modification Systems Associated with Niche Differentiation in the *Campylobacter jejuni* Sequence Type 403 Clonal Complex. *Appl Environ Microbiol* 81(11):3641–3647.

11. Mourkas E, et al. (2019) Gene pool transmission of multidrug resistance among *Campylobacter* from livestock, sewage and human disease. *Environ Microbiol* 00:1462-2920.14760.

12. Pascoe B, et al. (2017) Local genes for local bacteria: Evidence of allopatry in the genomes of transatlantic *Campylobacter* populations. *Mol Ecol* 26(17):4497–4508.

13. Richards VP, Lefébure T, Pavinski Bitar PD, Stanhope MJ (2013) Comparative characterization of the virulence gene clusters (lipooligosaccharide [LOS] and capsular polysaccharide [CPS]) for *Campylobacter coli*, *Campylobacter jejuni* subsp. *jejuni* and related *Campylobacter* species. *Infect Genet Evol* 14(1):200–213.

14. Sheppard SK, et al. (2014) Cryptic ecology among host generalist *Campylobacter jejuni* in

domestic animals. *Mol Ecol* 23(10):2442–2451.

15. Sheppard SK, et al. (2013) Progressive genome-wide introgression in agricultural *Campylobacter coli*. *Mol Ecol* 22(4):1051–1064.

16. Sheppard SK, et al. (2013) Genome-wide association study identifies vitamin B5 biosynthesis as a host specificity factor in *Campylobacter*. *Proc Natl Acad Sci* 110(29):11923–11927.

17. Takamiya M, et al. (2011) Genome Sequences of Two Stress-Tolerant *Campylobacter jejuni* Poultry Strains, 305 and DFVF1099. *J Bacteriol* 193(19):5546–5547.

18. Takamiya M, et al. (2011) Genome Sequence of *Campylobacter jejuni* strain 327, a strain isolated from a turkey slaughterhouse. *Stand Genomic Sci* 4(2):113–122.

19. Teh AHT, Lee SM, Dykes GA (2016) Draft Genome Sequences of Three Multiantibiotic-Resistant *Campylobacter jejuni* Strains (2865, 2868, and 2871) Isolated from Poultry at Retail Outlets in Malaysia. *Genome Announc* 4(3):e00331-16.

20. Thépault A, et al. (2017) Genome-Wide Identification of Host-Segregating Epidemiological Markers for Source Attribution in *Campylobacter jejuni*. *Appl Environ Microbiol* 83(7):e03085-16.

21. Yahara K, et al. (2016) Genome-wide association of functional traits linked with *Campylobacter jejuni* survival from farm to fork. *Environ Microbiol*:2–41.

22. Zeng X, Mo Y, Xu F, Lin J (2013) Identification and characterization of a periplasmic trilactone esterase, Cee, revealed unique features of ferric enterobactin acquisition in *Campylobacter*. *Mol Microbiol* 87(3):594–608.

23. Jolley KA, Bray JE, Maiden MCJ (2018) Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res* 3(0):124.

24. Taylor AJ, Zakai SAI, Kelly DJ (2017) The Periplasmic Chaperone Network of *Campylobacter jejuni*: Evidence that SalC (Cj1289) and PpiD (Cj0694) Are Involved in Maintaining Outer Membrane Integrity. *Front Microbiol* 8(MAR):1–14.

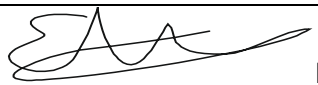
Appendix for chapter 4

4.1. Supplementary form SF3

4.2. Supplementary figures

4.3. Supplementary tables

4.1. Supplementary form SF3.

This declaration concerns the article entitled:									
Gene pool transmission of multidrug resistance among *Campylobacter* from livestock, sewage and human disease									
Publication status (tick one)									
Draft manuscript		**Submitted**		**In review**		**Accepted**		**Published**	X
Publication details (reference)	Mourkas, E., Florez-Cuadrado, D., Pascoe, B., Calland, J.K., Bayliss, S.C., Mageiros, L., Méric, G., Hitchings, M.D., Quesada, A., Porrero, C., Ugarte-Ruiz, M., Gutierrez-Fernandez, J., Dominguez, L., Sheppard, S.K. 2019. Gene pool transmission of multidrug resistance among *Campylobacter* from livestock, sewage and human disease. *Environmental Microbiology* 21 (12) 4597-4613. doi:10.1111/1462-2920.14760.								
Copyright status (tick the appropriate statement)									
I hold the copyright for this material	X	Copyright is retained by the publisher, but I have been given permission to replicate the material here							
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	The candidate contributed to / considerably contributed to / predominantly executed the… Formulation of ideas: Considerably contributed to 60% Design of methodology: Considerably contributed to 70% Experimental work: Considerably contributed to 60% Presentation of data in journal format: Considerably contributed to 70%								
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.								
Signed	 Evangelos Mourkas				**Date**	31/07/2020			

4.2. Supplementary figures

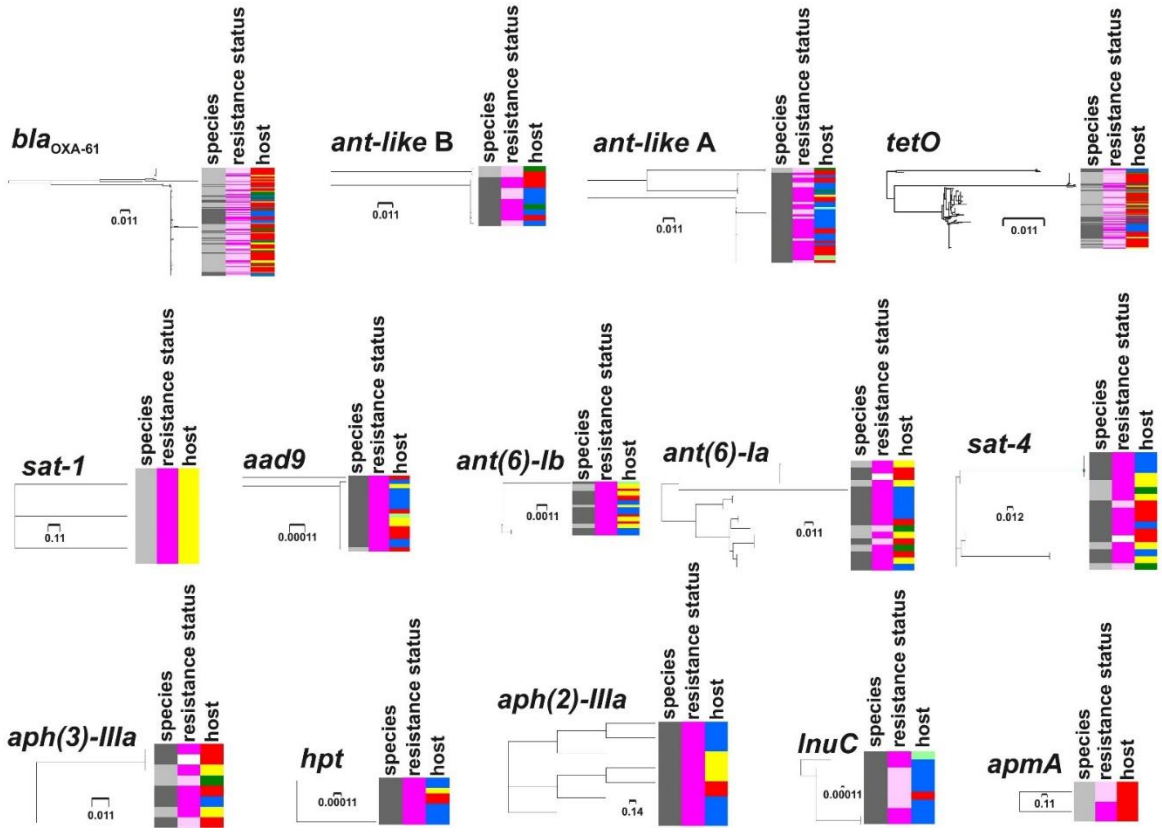


Figure S4.1. Individual AMR gene trees. 14 single-gene trees highlighting the allelic diversity in AMR genes found in *C. jejuni* (grey) and *C. coli* (black) isolates shown in the first column. The resistance status of each isolate is highlighted in the second column for multidrug resistant (dark pink), non-multidrug resistance (light pink) or not tested (white). The host of every isolate is shown in the third column for chickens (dark green), cattle (green), pigs (light green), humans (red) and sewage (blue). The scale represents the number of substitutions per site.

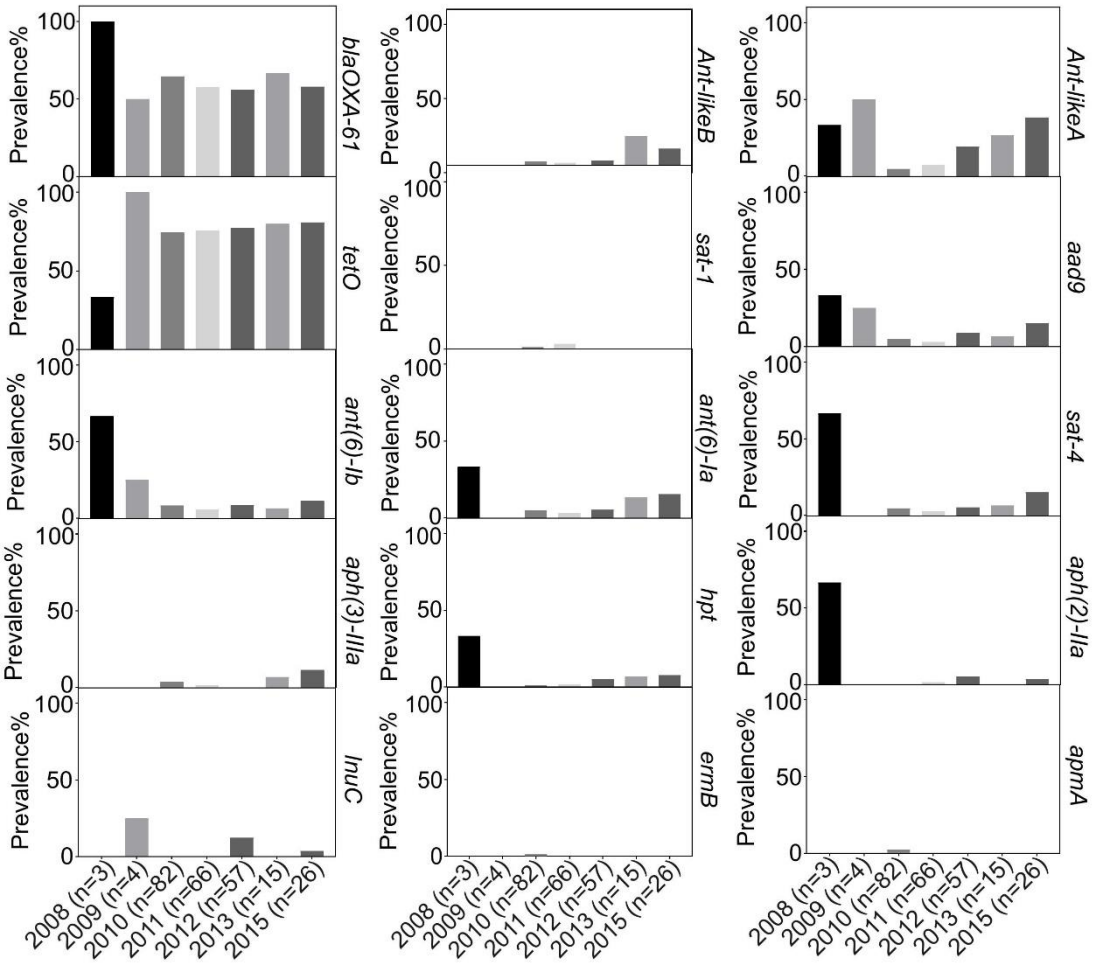


Figure S4.2. Prevalence of AMR genes over time. Graphs illustrate the presence of 15 putative AMR genes in isolate genomes sampled at each year in the chapter. Prevalence (%) was calculated by dividing the number of samples that had the AMR gene by the total number of samples in that year.

4.3. Supplementary tables

All supplementary tables are available at the public data repository FigShare (doi.org/10.6084/m9.figshare.12743327).

Table S4.1. Details of isolates used in chapter 4.

Table S4.2. Isolates and their MIC against different antibiotics used in chapter 4.

Table S4.3. Resistance phenotype-genotype correlations among *Campylobacter* isolates.

Table S4.4. Antibiotic drug classes: mechanism of action/resistance and AMR genes.

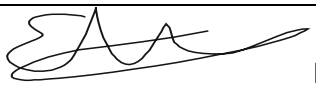
Table S4.5. Genomic and phenotypic details of all isolates used in chapter 4.

Appendix for chapter 5

5.1. Supplementary form SF4

5.2. Supplementary tables

5.1. Supplementary form SF4

This declaration concerns the article entitled:									
The potential of isolation source to predict colonization in avian hosts: a case study in *Campylobacter jejuni* strains from three bird species									
Publication status (tick one)									
Draft manuscript		**Submitted**		**In review**		**Accepted**		**Published**	X
Publication details (reference)	Atterby, C., Mourkas, E., Méric, G., Pascoe, B., Waldenström, J., Sheppard, S.K., Olsen, B., Järhult, J.D., Ellström, P. 2018. The potential of isolation source to predict colonization in avian hosts: a case study in *Campylobacter jejuni* strains from three bird species. *Frontiers in Microbiology* 9:51. doi:10.3389/fmicb.2018.00591.								
Copyright status (tick the appropriate statement)									
I hold the copyright for this material	X	Copyright is retained by the publisher, but I have been given permission to replicate the material here							
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	The candidate contributed to / considerably contributed to / predominantly executed the… Formulation of ideas: Contributed to 20% Design of methodology: Contributed to 40% Experimental work: considerably contributed to 70% Presentation of data in journal format: considerably contributed to 70%								
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.								
Signed	 Evangelos Mourkas					**Date**	31/07/2020		

5.2. Supplementary tables

All supplementary tables are available at the public data repository FigShare (doi.org/10.6084/m9.figshare.12739991).

Table S5.1. Isolate details used in chapter 5.

Table S5.2. List of genes loci belonging to STs from mallards, song thrushes and chickens.

Table S5.3. Summary of faecal counts from both infection experiments.

Table S5.4. Summary of qPCR CFU transformed values for infection experiment 2.

Table S5.5. Standard curve generated for qPCR.

.

References

Acheson, D. and Allos, B.M., 2001. Campylobacter jejuni Infections: Update on Emerging Issues and Trends. *Clinical Infectious Diseases*, 32(8), pp.1201–1206.

Achtman, M., 1996. A surfeit of YATMs? *Journal of clinical microbiology*, 34(7), pp.1870–1870.

Achtman, M., 2008. Evolution, Population Structure, and Phylogeography of Genetically Monomorphic Bacterial Pathogens. *Annual Review of Microbiology*, 62(1), pp.53–70.

Agency, F.S., 2014. *A microbiological survey of Campylobacter contamination in fresh whole UK-produced chilled chickens at retail sale—interim report to cover quarters 1–3* [Online]. Food Standards Agency, London, United Kingdom.

Alam, M.T., Petit, R.A., Crispell, E.K., Thornton, T.A., Conneely, K.N., Jiang, Y., Satola, S.W. and Read, T.D., 2014. Dissecting Vancomycin-Intermediate Resistance in Staphylococcus aureus Using Genome-Wide Association. *Genome Biology and Evolution*, 6(5), pp.1174–1185.

Alexander, J., Benford, D., Cockburn, A., Cravedi, J., Dogliotti, E., Domenico, A. Di, Férnandez-cruz, M.L., Fürst, P., Fink-gremmels, J., Galli, C.L., Grandjean, P., Gzyl, J., Heinemeyer, G., Mutti, A., Schlatter, J., Leeuwen, R. Van and Peteghem, C. Van, 2009. Nitrite as undesirable substances in animal feed - Scientific Opinion of the Panel on Contaminants in the Food Chain. *EFSA Journal*, 7(4), pp.1–47.

Amour, C., Gratz, J., Mduma, E., Svensen, E., Rogawski, E.T., McGrath, M., Seidman, J.C., McCormick, B.J.J., Shrestha, S., Samie, A., Mahfuz, M., Qureshi, S., Hotwani, A., Babji, S., Trigoso, D.R., Lima, A.A.M., Bodhidatta, L., Bessong, P., Ahmed, T., Shakoor, S., Kang, G., Kosek, M., Guerrant, R.L., Lang, D., Gottlieb, M., Houpt, E.R., Platts-Mills, J.A. and Etiology, Risk Factors, and I. of E.I. and M. and the C. for C.H. and D.P. (MAL-E.N.I., 2016. Epidemiology and Impact of Campylobacter Infection in Children in 8 Low-Resource Settings: Results From the MAL-ED Study. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 63(9), pp.1171–1179.

Andreani, N.A., Hesse, E. and Vos, M., 2017. Prokaryote genome fluidity is dependent on effective population size. *The ISME Journal*, 11(7), pp.1719–

1721.

Atterby, C., Mourkas, E., Méric, G., Pascoe, B., Wang, H., Waldenström, J., Sheppard, S.K., Olsen, B., Järhult, J.D. and Ellström, P., 2018. The Potential of Isolation Source to Predict Colonization in Avian Hosts: A Case Study in Campylobacter jejuni Strains From Three Bird Species. *Frontiers in Microbiology*, 9(MAR).

Axelsson-Olsson, D., Ellström, P., Waldenström, J., Haemig, P.D., Brudin, L. and Olsen, B., 2007. Acanthamoeba-Campylobacter coculture as a novel method for enrichment of Campylobacter species. *Applied and environmental microbiology*, 73(21), pp.6864–9.

Axelsson-Olsson, D., Waldenström, J., Broman, T., Olsen, B. and Holmberg, M., 2005. Protozoan Acanthamoeba polyphaga as a Potential Reservoir for Campylobacter jejuni. *Applied and Environmental Microbiology*, 71(2), pp.987–992.

Bacon, D.J., Alm, R.A., Hu, L., Hickey, T.E., Ewing, C.P., Batchelor, R.A., Trust, T.J. and Guerry, P., 2002. DNA Sequence and Mutational Analyses of the pVir Plasmid of Campylobacter jejuni 81-176. *Infection and Immunity*, 70(11), pp.6242–6250.

Baker, K.S., Dallman, T.J., Field, N., Childs, T., Mitchell, H., Day, M., Weill, F.-X., Lefèvre, S., Tourdjman, M., Hughes, G., Jenkins, C. and Thomson, N., 2018. Horizontal antimicrobial resistance transfer drives epidemics of multiple Shigella species. *Nature communications*, 9(1), p.1462.

den Bakker, H.C., Cummings, C.A., Ferreira, V., Vatta, P., Orsi, R.H., Degoricija, L., Barker, M., Petrauskene, O., Furtado, M.R. and Wiedmann, M., 2010. Comparative genomics of the bacterial genus Listeria: Genome evolution is characterized by limited gene acquisition and limited gene loss. *BMC Genomics*, 11(1), p.688.

den Bakker, H.C., Didelot, X., Fortes, E.D., Nightingale, K. and Wiedmann, M., 2008. Lineage specific recombination rates and microevolution in Listeria monocytogenes. *BMC Evolutionary Biology*, 8(1), p.277.

Balcazar, J.L., 2014. Bacteriophages as Vehicles for Antibiotic Resistance Genes in the Environment. *PLoS Pathogens*, 10(7), p.e1004219.

Bales, M.E., Dannenberg, A.L., Brachman, P.S., Kaufmann, A.F., Klatsky, P.C. and Ashford, D.A., 2002. Epidemiologic Responses to Anthrax Outbreaks: A Review of Field Investigations, 1950–2001. *Emerging Infectious Diseases*,

8(10), pp.1163–1174.

Balouiri, M., Sadiki, M. and Ibnsouda, S.K., 2016. Methods for in vitro evaluating antimicrobial activity: A review. *Journal of Pharmaceutical Analysis*, 6(2), pp.71–79.

Batchelor, R.A., 2004. Nucleotide sequences and comparison of two large conjugative plasmids from different Campylobacter species. *Microbiology*, 150(10), pp.3507–3517.

Batz, M.B., Henke, E. and Kowalcyk, B., 2013. Long-Term Consequences of Foodborne Infections. *Infectious Disease Clinics of North America*, 27(3), pp.599–616.

Bäumler, A.J., Hargis, B.M. and Tsolis, R.M., 2000. Tracing the origins of Salmonella outbreaks. *Science (New York, N.Y.)*, 287(5450), pp.50–2.

van Belkum, A., Tassios, P.T., Dijkshoorn, L., Haeggman, S., Cookson, B., Fry, N.K., Fussing, V., Green, J., Feil, E., Gerner-Smidt, P., Brisse, S. and Struelens, M., 2007. Guidelines for the validation and application of typing methods for use in bacterial epidemiology. *Clinical Microbiology and Infection*, 13, pp.1–46.

Bennett, G.M. and Moran, N.A., 2013. Small, Smaller, Smallest: The Origins and Evolution of Ancient Dual Symbioses in a Phloem-Feeding Insect. *Genome Biology and Evolution*, 5(9), pp.1675–1688.

Beres, S.B., Carroll, R.K., Shea, P.R., Sitkiewicz, I., Martinez-Gutierrez, J.C., Low, D.E., McGeer, A., Willey, B.M., Green, K., Tyrrell, G.J., Goldman, T.D., Feldgarden, M., Birren, B.W., Fofanov, Y., Boos, J., Wheaton, W.D., Honisch, C. and Musser, J.M., 2010. Molecular complexity of successive bacterial epidemics deconvoluted by comparative pathogenomics. *Proceedings of the National Academy of Sciences*, 107(9), pp.4371–4376.

van Bergen, M.A.P., Linnane, S., van Putten, J.P.M. and Wagenaar, J.A., 2005. Global detection and identification of Campylobacter fetus subsp. venerealis. *Revue scientifique et technique (International Office of Epizootics)*, 24(3), pp.1017–26.

Berthenet, E., Yahara, K., Thorell, K., Pascoe, B., Meric, G., Mikhail, J.M., Engstrand, L., Enroth, H., Burette, A., Megraud, F., Varon, C., Atherton, J.C., Smith, S., Wilkinson, T.S., Hitchings, M.D., Falush, D. and Sheppard, S.K., 2018. A GWAS on Helicobacter pylori strains points to genetic variants associated with gastric cancer risk. *BMC Biology*, 16(1), p.84.

Biswas, D., 2000. Uptake pathways of clinical and healthy animal isolates of Campylobacter jejuni into INT-407 cells. *FEMS Immunology and Medical Microbiology*, 29(3), pp.203–211.

Black, R.E., Levine, M.M., Clements, M. Lou, Hughes, T.P. and Blaser, M.J., 1988. Experimental Campylobacter jejuni Infection in Humans. *Journal of Infectious Diseases*, 157(3), pp.472–479.

Boerlin, P. and Reid-Smith, R.J., 2008. Antimicrobial resistance: its emergence and transmission. *Animal Health Research Reviews*, 9(02), pp.115–126.

Bolles, M., Donaldson, E. and Baric, R., 2011. SARS-CoV and emergent coronaviruses: viral determinants of interspecies transmission. *Current Opinion in Virology*, 1(6), pp.624–634.

Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., Matschiner, M., Mendes, F.K., Müller, N.F., Ogilvie, H.A., du Plessis, L., Popinga, A., Rambaut, A., Rasmussen, D., Siveroni, I., Suchard, M.A., Wu, C.-H., Xie, D., Zhang, C., Stadler, T. and Drummond, A.J., 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLOS Computational Biology*, 15(4), p.e1006650.

Bromham, L. and Penny, D., 2003. The modern molecular clock . *Nature Reviews Genetics*, 4(3), pp.216–224.

Brosch, R., Gordon, S. V., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., Garnier, T., Gutierrez, C., Hewinson, G., Kremer, K., Parsons, L.M., Pym, A.S., Samper, S., van Soolingen, D. and Cole, S.T., 2002. A new evolutionary scenario for the Mycobacterium tuberculosis complex. *Proceedings of the National Academy of Sciences*, 99(6), pp.3684–3689.

Bush, K., Courvalin, P., Dantas, G., Davies, J., Eisenstein, B., Huovinen, P., Jacoby, G.A., Kishony, R., Kreiswirth, B.N., Kutter, E., Lerner, S.A., Levy, S., Lewis, K., Lomovskaya, O., Miller, J.H., Mobashery, S., Piddock, L.J. V., Projan, S., Thomas, C.M., Tomasz, A., Tulkens, P.M., Walsh, T.R., Watson, J.D., Witkowski, J., Witte, W., Wright, G., Yeh, P. and Zgurskaya, H.I., 2011. Tackling antibiotic resistance. *Nature Reviews Microbiology*, 9(12), pp.894–896.

Castanon, J.I.R., 2007. History of the Use of Antibiotic as Growth Promoters in European Poultry Feeds. *Poultry Science*, 86(11), pp.2466–2471.

Chen, J., Quiles-Puchalt, N., Chiang, Y.N., Bacigalupe, R., Fillol-Salom, A., Chee, M.S.J., Fitzgerald, J.R. and Penadés, J.R., 2018. Genome hypermobility by

lateral transduction. *Science*, 362(6411), pp.207–212.

Chen, P.E. and Shapiro, B.J., 2015. The advent of genome-wide association studies for bacteria. *Current Opinion in Microbiology*, 25, pp.17–24.

Choi, K.-H., 2009. Applications of Transposon-Based Gene Delivery System in Bacteria. *Journal of Microbiology and Biotechnology*, 19(3), pp.217–28.

Chon, J.-W., Kim, H., Kim, H.-S. and Seo, K.-H., 2013. Improvement of modified charcoal-cefoperazone-deoxycholate agar by addition of potassium clavulanate for detecting Campylobacter spp. in chicken carcass rinse. *International Journal of Food Microbiology*, 165(1), pp.7–10.

Chun, J., Grim, C.J., Hasan, N.A., Lee, J.H., Choi, S.Y., Haley, B.J., Taviani, E., Jeon, Y.-S., Kim, D.W., Lee, J.-H., Brettin, T.S., Bruce, D.C., Challacombe, J.F., Detter, J.C., Han, C.S., Munk, A.C., Chertkov, O., Meincke, L., Saunders, E., Walters, R.A., Huq, A., Nair, G.B. and Colwell, R.R., 2009. Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic Vibrio cholerae. *Proceedings of the National Academy of Sciences*, 106(36), pp.15442–15447.

Chusid, M.J., Wortmann, D.W. and Dunne, W.M., 1990. “Campylobacter upsaliensis” sepsis in a boy with acquired hypogammaglobulinemia. *Diagnostic Microbiology and Infectious Disease*, 13(5), pp.367–369.

Coates, A., 2019. The future of antibiotics lies in combination treatments. *Future Drug Discovery*, 1(1), p.FDD5.

Cody, A.J., Maiden, M.C.J., Strachan, N.J.C. and McCarthy, N.D., 2019. A systematic review of source attribution of human campylobacteriosis using multilocus sequence typing. *Eurosurveillance*, 24(43).

Coker, A.O., Isokpehi, R.D., Thomas, B.N., Amisu, K.O. and Obi, C.L., 2002. Human campylobacteriosis in developing countries. *Emerging infectious diseases*, 8(3), pp.237–44.

Coleman, M.L., 2006. Genomic Islands and the Ecology and Evolution of Prochlorococcus. *Science*, 311(5768), pp.1768–1770.

Colles, F.M., Ali, J.S., Sheppard, S.K., McCarthy, N.D. and Maiden, M.C.J., 2011. Campylobacter populations in wild and domesticated Mallard ducks (Anas platyrhynchos). *Environmental Microbiology Reports*, 3(5), pp.574–580.

Collins, C. and Didelot, X., 2018. A phylogenetic method to perform genome-wide association studies in microbes that accounts for population structure and recombination. *PLOS Computational Biology*, 14(2), p.e1005958.

Corry, J.E.L., Post, D.E., Colin, P. and Laisney, M.J., 1995. Chapter 9 Culture media for the isolation of campylobacters. In: J.E.L. Corry, G.D.W. Curtis and R.M.B.T.-P. in I.M. Baird, eds. *Culture Media for Food Microbiology*. Elsevier, pp.129–162.

Corvin, A., Craddock, N. and Sullivan, P.F., 2010. Genome-wide association studies: a primer. *Psychological Medicine*, 40(7), pp.1063–1077.

Crim, S.M., Griffin, P.M., Tauxe, R., Marder, E.P., Gilliss, D., Cronquist, A.B., Cartter, M., Tobin-D'Angelo, M., Blythe, D., Smith, K., Lathrop, S., Zansky, S., Cieslak, P.R., Dunn, J., Holt, K.G., Wolpert, B., Henao, O.L. and Centers for Disease Control and Prevention (CDC), 2015. Preliminary incidence and trends of infection with pathogens transmitted commonly through food - Foodborne Diseases Active Surveillance Network, 10 U.S. sites, 2006-2014. *MMWR. Morbidity and mortality weekly report*, 64(18), pp.495–9.

Croucher, N.J., Finkelstein, J.A., Pelton, S.I., Mitchell, P.K., Lee, G.M., Parkhill, J., Bentley, S.D., Hanage, W.P. and Lipsitch, M., 2013. Population genomics of post-vaccine changes in pneumococcal epidemiology. *Nature Genetics*, 45(6), pp.656–663.

Croucher, N.J., Page, A.J., Connor, T.R., Delaney, A.J., Keane, J.A., Bentley, S.D., Parkhill, J. and Harris, S.R., 2015. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Research*, 43(3), pp.e15–e15.

Cui, Y., Yang, X., Didelot, X., Guo, C., Li, D., Yan, Y., Zhang, Y., Yuan, Y., Yang, H., Wang, Jian, Wang, Jun, Song, Y., Zhou, D., Falush, D. and Yang, R., 2015. Epidemic Clones, Oceanic Gene Pools, and Eco-LD in the Free Living Marine Pathogen Vibrio parahaemolyticus. *Molecular Biology and Evolution*, 32(6), pp.1396–1410.

Datta, S., Mitra, S., Chattopadhyay, P., Som, T., Mukherjee, S. and Basu, S., 2017. Spread and exchange of bla NDM-1 in hospitalized neonates: role of mobilizable genetic elements. *European Journal of Clinical Microbiology & Infectious Diseases*, 36(2), pp.255–265.

Dearlove, B.L., Cody, A.J., Pascoe, B., Méric, G., Wilson, D.J. and Sheppard, S.K., 2016. Rapid host switching in generalist Campylobacter strains erodes the signal for tracing human infections. *The ISME Journal*, 10(3), pp.721–729.

Derbise, A., de Cespedes, G. and Solh, N. El, 1997. Nucleotide sequence of theStaphylococcus aureus transposon, Tn5405, carrying aminoglycosides

resistance genes. *Journal of Basic Microbiology*, 37(5), pp.379–384.

Derbise, A., Dyke, K.G.H. and El Solh, N., 1996. Characterization of aStaphylococcus aureusTransposon, Tn5405,Located within Tn5404and Carrying the Aminoglycoside Resistance Genes,aphA-3andaadE. *Plasmid*, 35(3), pp.174–188.

Didelot, X., Croucher, N.J., Bentley, S.D., Harris, S.R. and Wilson, D.J., 2018. Bayesian inference of ancestral dates on bacterial phylogenetic trees. *Nucleic Acids Research*, 46(22), pp.e134–e134.

Didelot, X., Eyre, D.W., Cule, M., Ip, C.L.C., Ansari, M., Griffiths, D., Vaughan, A., O'Connor, L., Golubchik, T., Batty, E.M., Piazza, P., Wilson, D.J., Bowden, R., Donnelly, P.J., Dingle, K.E., Wilcox, M., Walker, A., Crook, D.W., A Peto, T.E. and Harding, R.M., 2012. Microevolutionary analysis of Clostridium difficile genomes to investigate transmission. *Genome Biology*, 13(12), p.R118.

Didelot, X., Urwin, R., Maiden, M.C.J. and Falush, D., 2009. Genealogical typing of Neisseria meningitidis. *Microbiology*, 155(10), pp.3176–3186.

Didelot, X. and Wilson, D.J., 2015. ClonalFrameML: Efficient Inference of Recombination in Whole Bacterial Genomes. *PLOS Computational Biology*, 11(2), p.e1004041.

Dingle, K.E., Colles, F.M., Wareing, D.R., Ure, R., Fox, A.J., Bolton, F.E., Bootsma, H.J., Willems, R.J., Urwin, R. and Maiden, M.C., 2001. Multilocus sequence typing system for Campylobacter jejuni. *Journal of clinical microbiology*, 39(1), pp.14–23.

Dong, E., Du, H. and Gardner, L., 2020. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases*, 3099(20), pp.19–20.

Doolittle, W.F. and Booth, A., 2017. It's the song, not the singer: an exploration of holobiosis and evolutionary theory. *Biology & Philosophy*, 32(1), pp.5–24.

Doolittle, W.F. and Zhaxybayeva, O., 2009. On the origin of prokaryotic species. *Genome Research*, 19(5), pp.744–756.

van Doorn, P.A., Ruts, L. and Jacobs, B.C., 2008. Clinical features, pathogenesis, and treatment of Guillain-Barré syndrome. *The Lancet Neurology*, 7(10), pp.939–950.

Drews, G., 2000. The roots of microbiology and the influence of Ferdinand Cohn on microbiology of the 19th century. *FEMS Microbiology Reviews*, 24(3), pp.225–249.

Drummond, A.J., Suchard, M.A., Xie, D. and Rambaut, A., 2012. Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, 29(8), pp.1969–1973.

Earle, S.G., Wu, C.-H., Charlesworth, J., Stoesser, N., Gordon, N.C., Walker, T.M., Spencer, C.C.A., Iqbal, Z., Clifton, D.A., Hopkins, K.L., Woodford, N., Smith, E.G., Ismail, N., Llewelyn, M.J., Peto, T.E., Crook, D.W., McVean, G., Walker, A.S. and Wilson, D.J., 2016. Identifying lineage effects when controlling for population structure improves power in bacterial association studies. *Nature Microbiology*, 1(5), p.16041.

Egan, E.S., Fogel, M.A. and Waldor, M.K., 2005. MicroReview: Divided genomes: negotiating the cell cycle in prokaryotes with multiple chromosomes. *Molecular Microbiology*, 56(5), pp.1129–1138.

Eldholm, V., Monteserin, J., Rieux, A., Lopez, B., Sobkowiak, B., Ritacco, V. and Balloux, F., 2015. Four decades of transmission of a multidrug-resistant Mycobacterium tuberculosis outbreak strain. *Nature Communications*, 6(1), p.7119.

Ellis-Iversen, J., Cook, A.J.C., Smith, R.P., Pritchard, G.C. and Nielen, M., 2009. Temporal patterns and risk factors for Escherichia coli O157 and Campylobacter spp, in young cattle. *Journal of food protection*, 72(3), pp.490–6.

Ellis-Iversen, J., Ridley, A., Morris, V., Sowa, A., Harris, J., Atterbury, R., Sparks, N. and Allen, V., 2012. Persistent environmental reservoirs on farms as risk factors for Campylobacter in commercial poultry. *Epidemiology and infection*, 140(5), pp.916–24.

Engberg, J., Aarestrup, F.M., Taylor, D.E., Gerner-Smidt, P. and Nachamkin, I., 2001. Quinolone and macrolide resistance in Campylobacter jejuni and C. coli: resistance mechanisms and trends in human isolates. *Emerging infectious diseases*, 7(1), pp.24–34.

Enright, M.C. and Spratt, B.G., 1998. A multilocus sequence typing scheme for Streptococcus pneumoniae: identification of clones associated with serious invasive disease. *Microbiology*, 144(11), pp.3049–3060.

Enright, M.C. and Spratt, B.G., 1999. Multilocus sequence typing. *Trends in Microbiology*, 7(12), pp.482–487.

Eurosurveillance editorial team, 2014. The European Union Summary Report on Trends and Sources of Zoonoses, Zoonotic Agents and Food-borne

Outbreaks in 2012. *EFSA Journal*, 12(2), p.3547.

Everitt, R.G., Didelot, X., Batty, E.M., Miller, R.R., Knox, K., Young, B.C., Bowden, R., Auton, A., Votintseva, A., Larner-Svensson, H., Charlesworth, J., Golubchik, T., Ip, C.L.C., Godwin, H., Fung, R., Peto, T.E.A., Walker, A.S., Crook, D.W. and Wilson, D.J., 2014. Mobile elements drive recombination hotspots in the core genome of Staphylococcus aureus. *Nature Communications*, 5(1), p.3956.

Falush, D., Kraft, C., Taylor, N.S., Correa, P., Fox, J.G., Achtman, M. and Suerbaum, S., 2001. Recombination and mutation during long-term gastric colonization by Helicobacter pylori: Estimates of clock rates, recombination size, and minimal age. *Proceedings of the National Academy of Sciences*, 98(26), pp.15056–15061.

Falush, D., Wirth, T., Linz, B., Pritchard, K.J., Stephens, M., Kidd, M., Blaser, J.M., Graham, Y.D., Vacher, S., Perez-Perez, I.G., Yamaoka, Y., Megraud, F., Otto, K., Reichard, U., Katzowitsch, E., Wang, X., Achtman, M. and Suerbaum, S., 2003. Traces of Human Migrations in Helicobacter pylori Populations. *Science*, 299(5612), pp.1582–1585.

Farhat, M.R., Shapiro, B.J., Kieser, K.J., Sultana, R., Jacobson, K.R., Victor, T.C., Warren, R.M., Streicher, E.M., Calver, A., Sloutsky, A., Kaur, D., Posey, J.E., Plikaytis, B., Oggioni, M.R., Gardy, J.L., Johnston, J.C., Rodrigues, M., Tang, P.K.C., Kato-Maeda, M., Borowsky, M.L., Muddukrishna, B., Kreiswirth, B.N., Kurepina, N., Galagan, J., Gagneux, S., Birren, B., Rubin, E.J., Lander, E.S., Sabeti, P.C. and Murray, M., 2013. Genomic analysis identifies targets of convergent positive selection in drug-resistant Mycobacterium tuberculosis. *Nature Genetics*, 45(10), pp.1183–1189.

Faria, N.R., Rambaut, A., Suchard, M.A., Baele, G., Bedford, T., Ward, M.J., Tatem, A.J., Sousa, J.D., Arinaminpathy, N., Pépin, J., Posada, D., Peeters, M., Pybus, O.G. and Lemey, P., 2014. The early spread and epidemic ignition of HIV-1 in human populations. *Science*, 346(6205), pp.56–61.

Feil, E.J., Cooper, J.E., Grundmann, H., Robinson, D.A., Enright, M.C., Berendt, T., Peacock, S.J., Smith, J.M., Murphy, M., Spratt, B.G., Moore, C.E. and Day, N.P.J., 2003. How Clonal Is Staphylococcus aureus? *Journal of Bacteriology*, 185(11), pp.3307–3316.

Fitzgerald, C., Tu, Z. chao, Patrick, M., Stiles, T., Lawson, A.J., Santovenia, M., Gilbert, M.J., van Bergen, M., Joyce, K., Pruckler, J., Stroika, S., Duim, B.,

Miller, W.G., Loparev, V., Sinnige, J.C., Fields, P.I., Tauxe, R. V., Blaser, M.J. and Wagenaar, J.A., 2014. Campylobacter fetus subsp. testudinum subsp. nov., isolated from humans and reptiles. *International Journal of Systematic and Evolutionary Microbiology*, 64(Pt_9), pp.2944–2948.

Fleischmann, R., Adams, M., White, O., Clayton, R., Kirkness, E., Kerlavage, A., Bult, C., Tomb, J., Dougherty, B., Merrick, J. and Al., E., 1995. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science*, 269(5223), pp.496–512.

Flemming, H.-C. and Wuertz, S., 2019. Bacteria and archaea on Earth and their abundance in biofilms. *Nature reviews. Microbiology*, 17(4), pp.247–260.

Florez-Cuadrado, D., Ugarte-Ruiz, M., Meric, G., Quesada, A., Porrero, M.C., Pascoe, B., Sáez-Llorente, J.L., Orozco, G.L., Domínguez, L. and Sheppard, S.K., 2017. Genome Comparison of Erythromycin Resistant Campylobacter from Turkeys Identifies Hosts and Pathways for Horizontal Spread of erm(B) Genes. *Frontiers in Microbiology*, 8(NOV), pp.1–8.

Florez-Cuadrado, D., Ugarte-Ruiz, M., Quesada, A., Palomo, G., Domínguez, L. and Porrero, M.C., 2016. Description of an erm (B)-carrying Campylobacter coli isolate in Europe. *Journal of Antimicrobial Chemotherapy*, 71(3), pp.841–843.

Food, E. and Authority, S., 2019. The European Union summary report on antimicrobial resistance in zoonotic and indicator bacteria from humans, animals and food in 2017. *EFSA Journal*, 17(2).

Fox, J.G., Ackerman, J.I., Taylor, N., Claps, M. and Murphy, J.C., 1987. Campylobacter jejuni infection in the ferret: an animal model of human campylobacteriosis. *American journal of veterinary research*, 48(1), pp.85–90.

Fox, J.G., Rogers, A.B., Whary, M.T., Ge, Z., Taylor, N.S., Xu, S., Horwitz, B.H. and Erdman, S.E., 2004. Gastroenteritis in NF-κB-Deficient Mice Is Produced with Wild-Type Camplyobacter jejuni but Not with C. jejuni Lacking Cytolethal Distending Toxin despite Persistent Colonization with Both Strains. *Infection and Immunity*, 72(2), pp.1116–1125.

Frank, C., Werber, D., Cramer, J.P., Askar, M., Faber, M., an der Heiden, M., Bernard, H., Fruth, A., Prager, R., Spode, A., Wadl, M., Zoufaly, A., Jordan, S., Kemper, M.J., Follin, P., Müller, L., King, L.A., Rosner, B., Buchholz, U., Stark, K. and Krause, G., 2011. Epidemic Profile of Shiga-Toxin–Producing Escherichia coli O104:H4 Outbreak in Germany. *New England Journal of*

Medicine, 365(19), pp.1771–1780.

Fravalo, P., Laisney, M.-J., Gillard, M.-O., Salvat, G. and Chemaly, M., 2009. Campylobacter transfer from naturally contaminated chicken thighs to cutting boards is inversely related to initial load. *Journal of food protection*, 72(9), pp.1836–40.

Freddo, L., Yu, R., Latov, N., Donofrio, P., Hays, A., Greenberg, H., Albers, J., Allessi, A. and Keren, D., 1986. Gangliosides GM1 and GD1b are antigens for IgM M-protein in a patient with motor neuron disease. *Neurology*, 36(4), pp.454–458.

Fry, B.N., Feng, S., Chen, Y.Y., Newell, D.G., Coloe, P.J. and Korolik, V., 2000. The galE gene of Campylobacter jejuni is involved in lipopolysaccharide synthesis and virulence. *Infection and immunity*, 68(5), pp.2594–601.

Gagneux, S., 2018. Ecology and evolution of Mycobacterium tuberculosis. *Nature Reviews Microbiology*, 16(4), pp.202–213.

Gautret, P., Cramer, J.P., Field, V., Caumes, E., Jensenius, M., Gkrania-Klotsas, E., de Vries, P.J., Grobusch, M.P., Lopez-Velez, R., Castelli, F., Schlagenhauf, P., Hervius Askling, H., von Sonnenburg, F., Lalloo, D.G., Loutan, L., Rapp, C., Basto, F., Santos O'Connor, F., Weld, L., Parola, P. and EuroTravNet Network, 2012. Infectious diseases among travellers and migrants in Europe, EuroTravNet 2010. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin*, 17(26).

Giacomelli, M., Andrighetto, C., Lombardi, A., Martini, M. and Piccirillo, A., 2012. A Longitudinal Study on Thermophilic Campylobacter spp. in Commercial Turkey Flocks in Northern Italy: Occurrence and Genetic Diversity. *Avian Diseases*, 56(4), pp.693–700.

Giacomelli, M. and Piccirillo, A., 2014. Pet reptiles as potential reservoir of Campylobacter species with zoonotic potential: TABLE 1: *Veterinary Record*, 174(19), pp.479.2-479.

Gibreel, A., 2006. Macrolide resistance in Campylobacter jejuni and Campylobacter coli. *Journal of Antimicrobial Chemotherapy*, 58(2), pp.243–255.

Gilbert, M., Karwaski, M.-F., Bernatchez, S., Young, N.M., Taboada, E., Michniewicz, J., Cunningham, A.-M. and Wakarchuk, W.W., 2002. The genetic bases for the variation in the lipo-oligosaccharide of the mucosal

pathogen, Campylobacter jejuni. Biosynthesis of sialylated ganglioside mimics in the core oligosaccharide. *The Journal of biological chemistry*, 277(1), pp.327–37.

Girard, M.P., Tam, J.S., Assossou, O.M. and Kieny, M.P., 2010. The 2009 A (H1N1) influenza virus pandemic: A review. *Vaccine*, 28(31), pp.4895–4902.

Gire, S.K., Goba, A., Andersen, K.G., Sealfon, R.S.G., Park, D.J., Kanneh, L., Jalloh, S., Momoh, M., Fullah, M., Dudas, G., Wohl, S., Moses, L.M., Yozwiak, N.L., Winnicki, S., Matranga, C.B., Malboeuf, C.M., Qu, J., Gladden, A.D., Schaffner, S.F., Yang, X., Jiang, P.-P., Nekoui, M., Colubri, A., Coomber, M.R., Fonnie, M., Moigboi, A., Gbakie, M., Kamara, F.K., Tucker, V., Konuwa, E., Saffa, S., Sellu, J., Jalloh, A.A., Kovoma, A., Koninga, J., Mustapha, I., Kargbo, K., Foday, M., Yillah, M., Kanneh, F., Robert, W., Massally, J.L.B., Chapman, S.B., Bochicchio, J., Murphy, C., Nusbaum, C., Young, S., Birren, B.W., Grant, D.S., Scheiffelin, J.S., Lander, E.S., Happi, C., Gevao, S.M., Gnirke, A., Rambaut, A., Garry, R.F., Khan, S.H. and Sabeti, P.C., 2014. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. *Science*, 345(6202), pp.1369–1372.

Giske, C.G., Fröding, I., Hasan, C.M., Turlej-Rogacka, A., Toleman, M., Livermore, D., Woodford, N. and Walsh, T.R., 2012. Diverse Sequence Types of Klebsiella pneumoniae Contribute to the Dissemination of bla NDM-1 in India, Sweden, and the United Kingdom. *Antimicrobial Agents and Chemotherapy*, 56(5), pp.2735–2738.

Gradel, K.O., Nielsen, H.L., Schønheyder, H.C., Ejlertsen, T., Kristensen, B. and Nielsen, H., 2009. Increased Short- and Long-Term Risk of Inflammatory Bowel Disease After Salmonella or Campylobacter Gastroenteritis. *Gastroenterology*, 137(2), pp.495–501.

Griekspoor, P., Colles, F.M., McCarthy, N.D., Hansbro, P.M., Ashhurst-Smith, C., Olsen, B., Hasselquist, D., Maiden, M.C.J. and Waldenström, J., 2013. Marked host specificity and lack of phylogeographic population structure of Campylobacter jejuni in wild birds. *Molecular Ecology*, 22(5), pp.1463–1472.

Griekspoor, P., Olsen, B. and Waldenström, J., 2009. Campylobacter jejuni in penguins, Antarctica. *Emerging infectious diseases*, 15(5), pp.847–8.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W. and Gascuel, O., 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology*,

59(3), pp.307–321.

Guyard-Nicodème, M., Tresse, O., Houard, E., Jugiau, F., Courtillon, C., El Manaa, K., Laisney, M.-J. and Chemaly, M., 2013. Characterization of Campylobacter spp. transferred from naturally contaminated chicken legs to cooked chicken slices via a cutting board. *International Journal of Food Microbiology*, 164(1), pp.7–14.

Hacker, J., Blum-Oehler, G., Muhldorfer, I. and Tschape, H., 1997. Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Molecular Microbiology*, 23(6), pp.1089–1097.

Hamilton, H.L. and Dillard, J.P., 2006. Natural transformation of Neisseria gonorrhoeae : from DNA donation to homologous recombination. *Molecular Microbiology*, 59(2), pp.376–385.

Han, K., Li, Z., Peng, R., Zhu, L., Zhou, T., Wang, L., Li, S., Zhang, X., Hu, W., Wu, Z., Qin, N. and Li, Y., 2013. Extraordinary expansion of a Sorangium cellulosum genome from an alkaline milieu. *Scientific Reports*, 3(1), p.2101.

Hanage, W.P., Fraser, C., Tang, J., Connor, T.R. and Corander, J., 2009. Hyper-Recombination, Diversity, and Antibiotic Resistance in Pneumococcus. *Science*, 324(5933), pp.1454–1457.

Hänninen, M.-L., Haajanen, H., Pummi, T., Wermundsen, K., Katila, M.-L., Sarkkinen, H., Miettinen, I. and Rautelin, H., 2003. Detection and Typing of Campylobacter jejuni and Campylobacter coli and Analysis of Indicator Organisms in Three Waterborne Outbreaks in Finland. *Applied and Environmental Microbiology*, 69(3), pp.1391–1396.

Hara-Kudo, Y. and Takatori, K., 2011. Contamination level and ingestion dose of foodborne pathogens associated with infections. *Epidemiology and infection*, 139(10), pp.1505–10.

Hatfull, G.F., 2008. Bacteriophage genomics. *Current Opinion in Microbiology*, 11(5), pp.447–453.

Havelaar, A.H., Ivarsson, S., Löfdahl, M. and Nauta, M.J., 2013. Estimating the true incidence of campylobacteriosis and salmonellosis in the European Union, 2009. *Epidemiology and infection*, 141(2), pp.293–302.

Hermans, D., Pasmans, F., Messens, W., Martel, A., Van Immerseel, F., Rasschaert, G., Heyndrickx, M., Van Deun, K. and Haesebrouck, F., 2012. Poultry as a Host for the Zoonotic Pathogen Campylobacter jejuni. *Vector-Borne and Zoonotic Diseases*, 12(2), pp.89–98.

Hickey, T.E., McVeigh, A.L., Scott, D.A., Michielutti, R.E., Bixby, A., Carroll, S.A., Bourgeois, A.L. and Guerry, P., 2000. Campylobacter jejuni Cytolethal Distending Toxin Mediates Release of Interleukin-8 from Intestinal Epithelial Cells. *Infection and Immunity*, 68(12), pp.6535–6541.

Hormeño, L., Palomo, G., Ugarte-Ruiz, M., Porrero, M.C., Borge, C., Vadillo, S., Píriz, S., Domínguez, L., Campos, M.J. and Quesada, A., 2016. Identification of the main quinolone resistance determinant in Campylobacter jejuni and Campylobacter coli by MAMA-DEG PCR. *Diagnostic Microbiology and Infectious Disease*, 84(3), pp.236–239.

Hottes, A.K., Freddolino, P.L., Khare, A., Donnell, Z.N., Liu, J.C. and Tavazoie, S., 2013. Bacterial Adaptation through Loss of Function. *PLoS Genetics*, 9(7), p.e1003617.

Hug, L.A., Baker, B.J., Anantharaman, K., Brown, C.T., Probst, A.J., Castelle, C.J., Butterfield, C.N., Hernsdorf, A.W., Amano, Y., Ise, K., Suzuki, Y., Dudek, N., Relman, D.A., Finstad, K.M., Amundson, R., Thomas, B.C. and Banfield, J.F., 2016. A new view of the tree of life. *Nature Microbiology*, 1(5), p.16048.

Hug, L.A., Castelle, C.J., Wrighton, K.C., Thomas, B.C., Sharon, I., Frischkorn, K.R., Williams, K.H., Tringe, S.G. and Banfield, J.F., 2013. Community genomic analyses constrain the distribution of metabolic traits across the Chloroflexi phylum and indicate roles in sediment carbon cycling. *Microbiome*, 1(1), p.22.

Huttner, A., Harbarth, S., Carlet, J., Cosgrove, S., Goossens, H., Holmes, A., Jarlier, V., Voss, A. and Pittet, D., 2013. Antimicrobial resistance: a global view from the 2013 World Healthcare-Associated Infections Forum. *Antimicrobial Resistance and Infection Control*, 2(1), p.31.

Iraola, G., Forster, S.C., Kumar, N., Lehours, P., Bekal, S., García-Peña, F.J., Paolicchi, F., Morsella, C., Hotzel, H., Hsueh, P.-R., Vidal, A., Lévesque, S., Yamazaki, W., Balzan, C., Vargas, A., Piccirillo, A., Chaban, B., Hill, J.E., Betancor, L., Collado, L., Truyers, I., Midwinter, A.C., Dagi, H.T., Mégraud, F., Calleros, L., Pérez, R., Naya, H. and Lawley, T.D., 2017. Distinct Campylobacter fetus lineages adapted as livestock pathogens and human pathobionts in the intestinal microbiota. *Nature Communications*, 8(1), p.1367.

Jackson, B.R., Zegarra, J.A., López-Gatell, H., Sejvar, J., Arzate, F., Waterman, S., Núñez, A.S., López, B., Weiss, J., Cruz, R.Q., Murrieta, D.Y.L., Luna-Gierke, R., Heiman, K., Vieira, A.R., Fitzgerald, C., Kwan, P., Zárate-

Bermúdez, M., Talkington, D., Hill, V.R., Mahon, B. and GBS Outbreak Investigation Team, 2014. Binational outbreak of Guillain-Barré syndrome associated with Campylobacter jejuni infection, Mexico and USA, 2011. *Epidemiology and infection*, 142(5), pp.1089–99.

Jaillard, M., Lima, L., Tournoud, M., Mahé, P., van Belkum, A., Lacroix, V. and Jacob, L., 2018. A fast and agnostic method for bacterial genome-wide association studies: Bridging the gap between k-mers and genetic events. *PLOS Genetics*, 14(11), p.e1007758.

Janssen, R., Krogfelt, K.A., Cawthraw, S.A., van Pelt, W., Wagenaar, J.A. and Owen, R.J., 2008. Host-Pathogen Interactions in Campylobacter Infections: the Host Perspective. *Clinical Microbiology Reviews*, 21(3), pp.505–518.

Jasson, V., Sampers, I., Botteldoorn, N., López-Gálvez, F., Baert, L., Denayer, S., Rajkovic, A., Habib, I., De Zutter, L. and Debevere, J., 2009. Characterization of Escherichia coli from raw poultry in Belgium and impact on the detection of Campylobacter jejuni using Bolton broth. *International Journal of Food Microbiology*, 135(3), pp.248–253.

Jolley, Keith A., Bliss, C.M., Bennett, J.S., Bratcher, H.B., Brehony, C., Colles, F.M., Wimalarathna, H., Harrison, O.B., Sheppard, S.K., Cody, A.J. and Maiden, M.C.J., 2012. Ribosomal multilocus sequence typing: universal characterization of bacteria from domain to strain. *Microbiology*, 158(4), pp.1005–1015.

Jolley, Keith A, Hill, D.M.C., Bratcher, H.B., Harrison, O.B., Feavers, I.M., Parkhill, J. and Maiden, M.C.J., 2012. Resolution of a Meningococcal Disease Outbreak from Whole-Genome Sequence Data with Rapid Web-Based Analysis Methods. *Journal of Clinical Microbiology*, 50(9), pp.3046–3053.

Jolley, K.A. and Maiden, M.C.J., 2010. BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC bioinformatics*, 11(1), p.595.

Kaakoush, N.O., Castaño-Rodríguez, N., Mitchell, H.M. and Man, S.M., 2015. Global Epidemiology of Campylobacter Infection. *Clinical microbiology reviews*, 28(3), pp.687–720.

Karlyshev, A. V., Linton, D., Gregson, N.A. and Wren, B.W., 2002. A novel paralogous gene family involved in phase-variable flagella-mediated motility in Campylobacter jejuni. *Microbiology*, 148(2), pp.473–480.

Kettler, G.C., Martiny, A.C., Huang, K., Zucker, J., Coleman, M.L., Rodrigue, S., Chen, F., Lapidus, A., Ferriera, S., Johnson, J., Steglich, C., Church, G.M.,

Richardson, P. and Chisholm, S.W., 2007. Patterns and Implications of Gene Gain and Loss in the Evolution of Prochlorococcus. *PLoS Genetics*, 3(12), p.e231.

Kim, S.A., Lee, Y.M., Hwang, I.G., Kang, D.H., Woo, G.J. and Rhee, M.S., 2009. Eight enrichment broths for the isolation of Campylobacter jejuni from inoculated suspensions and ground pork. *Letters in Applied Microbiology*, 49(5), pp.620–626.

King, D.A., Peckham, C., Waage, J.K., Brownlie, J. and Woolhouse, M.E.J., 2006. Infectious Diseases: Preparing for the Future. *Science*, 313(5792), pp.1392–1393.

Kirk, K.F., Méric, G., Nielsen, H.L., Pascoe, B., Sheppard, S.K., Thorlacius-Ussing, O. and Nielsen, H., 2018. Molecular epidemiology and comparative genomics of Campylobacter concisus strains from saliva, faeces and gut mucosal biopsies in inflammatory bowel disease. *Scientific Reports*, 8(1), p.1902.

Klein, R.J., Zeiss, C., Chew, E.Y., Tsai, J.-Y., Sackler, R.S., Haynes, C., Henning, A.K., SanGiovanni, J.P., Mane, S.M., Mayne, S.T., Bracken, M.B., Ferris, F.L., Ott, J., Barnstable, C. and Hoh, J., 2005. Complement factor H polymorphism in age-related macular degeneration. *Science (New York, N.Y.)*, 308(5720), pp.385–9.

Kohler, R., Krause, G., Beutin, L., Stephan, R. and Zweifel, C., 2008. Shedding of food-borne pathogens and microbiological carcass contamination in rabbits at slaughter. *Veterinary Microbiology*, 132(1–2), pp.149–157.

Koskiniemi, S., Sun, S., Berg, O.G. and Andersson, D.I., 2012. Selection-Driven Gene Loss in Bacteria. *PLoS Genetics*, 8(6), p.e1002787.

Kulkarni, S.P., Lever, S., Logan, J.M.J., Lawson, A.J., Stanley, J. and Shafi, M.S., 2002. Detection of campylobacter species: a comparison of culture and polymerase chain reaction based methods. *Journal of clinical pathology*, 55(10), pp.749–53.

Kuo, C.-H. and Ochman, H., 2009. Inferring clocks when lacking rocks: the variable rates of molecular evolution in bacteria. *Biology direct*, 4(1), p.35.

de la Cruz, F. and Davies, J., 2000. Horizontal gene transfer and the origin of species: lessons from bacteria. *Trends in Microbiology*, 8(3), pp.128–133.

Laabei, M., Recker, M., Rudkin, J.K., Aldeljawi, M., Gulay, Z., Sloan, T.J., Williams, P., Endres, J.L., Bayles, K.W., Fey, P.D., Yajjala, V.K., Widhelm, T., Hawkins, E., Lewis, K., Parfett, S., Scowen, L., Peacock, S.J., Holden, M.,

Wilson, D., Read, T.D., van den Elsen, J., Priest, N.K., Feil, E.J., Hurst, L.D., Josefsson, E. and Massey, R.C., 2014. Predicting the virulence of MRSA from its genome sequence. *Genome Research*, 24(5), pp.839–849.

Land, M., Hauser, L., Jun, S.-R., Nookaew, I., Leuze, M.R., Ahn, T.-H., Karpinets, T., Lund, O., Kora, G., Wassenaar, T., Poudel, S. and Ussery, D.W., 2015. Insights from 20 years of bacterial genome sequencing. *Functional & Integrative Genomics*, 15(2), pp.141–161.

Lapierre, L., Arias, M.L. and Fernández, H., 2016. Antimicrobial Resistance in Campylobacter spp. *Campylobacter spp. and Related Organisms in Poultry*. Cham: Springer International Publishing, pp.165–183.

Lawrence, J.G., 1999. Gene transfer, speciation, and the evolution of bacterial genomes. *Current Opinion in Microbiology*, 2(5), pp.519–523.

Leatherbarrow, A.J.H., Griffiths, R., Hart, C.A., Kemp, R., Williams, N.J., Diggle, P.J., Wright, E.J., Sutherst, J., Houghton, P. and French, N.P., 2007. Campylobacter lari : genotype and antibiotic resistance of isolates from cattle, wildlife and water in an area of mixed dairy farmland in the United Kingdom. *Environmental Microbiology*, 9(7), pp.1772–1779.

Lee, G., Pan, W., Peñataro Yori, P., Paredes Olortegui, M., Tilley, D., Gregory, M., Oberhelman, R., Burga, R., Chavez, C.B. and Kosek, M., 2013. Symptomatic and Asymptomatic Campylobacter Infections Associated with Reduced Growth in Peruvian Children. *PLoS Neglected Tropical Diseases*, 7(1), p.e2036.

Lees, J.A., Vehkala, M., Välimäki, N., Harris, S.R., Chewapreecha, C., Croucher, N.J., Marttinen, P., Davies, M.R., Steer, A.C., Tong, S.Y.C., Honkela, A., Parkhill, J., Bentley, S.D. and Corander, J., 2016. Sequence element enrichment analysis to determine the genetic basis of bacterial phenotypes. *Nature Communications*, 7(1), p.12797.

Lenski, R.E., 2017. Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. *The ISME Journal*, 11(10), pp.2181–2194.

Leski, T.A., Bangura, U., Jimmy, D.H., Ansumana, R., Lizewski, S.E., Stenger, D.A., Taitt, C.R. and Vora, G.J., 2013. Multidrug-resistant tet(X)-containing hospital isolates in Sierra Leone. *International Journal of Antimicrobial Agents*, 42(1), pp.83–86.

Levesque, S., Frost, E., Arbeit, R.D. and Michaud, S., 2008. Multilocus Sequence

Typing of Campylobacter jejuni Isolates from Humans, Chickens, Raw Milk, and Environmental Water in Quebec, Canada. *Journal of Clinical Microbiology*, 46(10), pp.3404–3411.

Li, M., Du, X., Villaruz, A.E., Diep, B.A., Wang, D., Song, Y., Tian, Y., Hu, J., Yu, F., Lu, Y. and Otto, M., 2012. MRSA epidemic linked to a quickly spreading colonization and virulence determinant. *Nature Medicine*, 18(5), pp.816–819.

Livermore, D.M., 2007. Introduction: the challenge of multiresistance. *International Journal of Antimicrobial Agents*, 29(SUPPL. 3), pp.S1–S7.

Llosa, M., Gomis-Ruth, F.X., Coll, M. and Cruz, F. de la, 2002. Bacterial conjugation: a two-step mechanism for DNA transport. *Molecular Microbiology*, 45(1), pp.1–8.

Locey, K.J. and Lennon, J.T., 2016. Scaling laws predict global microbial diversity. *Proceedings of the National Academy of Sciences*, 113(21), pp.5970–5975.

Loman, N.J. and Pallen, M.J., 2015. Twenty years of bacterial genome sequencing. *Nature Reviews Microbiology*, 13(12), pp.787–794.

Lowder, B. V., Guinane, C.M., Ben Zakour, N.L., Weinert, L.A., Conway-Morris, A., Cartwright, R.A., Simpson, A.J., Rambaut, A., Nubel, U. and Fitzgerald, J.R., 2009. Recent human-to-poultry host jump, adaptation, and pandemic spread of Staphylococcus aureus. *Proceedings of the National Academy of Sciences*, 106(46), pp.19545–19550.

Luangtongkum, T., Jeon, B., Han, J., Plummer, P., Logue, C.M. and Zhang, Q., 2009. Antibiotic resistance in Campylobacter : emergence, transmission and persistence. *Future Microbiology*, 4(2), pp.189–200.

Luo, N., Sahin, O., Lin, J., Michel, L.O. and Zhang, Q., 2003. In Vivo Selection of Campylobacter Isolates with High Levels of Fluoroquinolone Resistance Associated with gyrA Mutations and the Function of the CmeABC Efflux Pump. *Antimicrobial Agents and Chemotherapy*, 47(1), pp.390–394.

Macfarlane, L., Walker, J., Borrow, R., Oppenheim, B.A. and Fox, A.J., 1999. Improved recognition of MRSA case clusters by the application of molecular subtyping using pulsed-field gel electrophoresis. *Journal of Hospital Infection*, 41(1), pp.29–37.

Maiden, M.C.J., Bygraves, J.A., Feil, E., Morelli, G., Russell, J.E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D.A., Feavers, I.M., Achtman, M. and Spratt, B.G., 1998. Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms.

Proceedings of the National Academy of Sciences, 95(6), pp.3140–3145.

Maiden, M.C.J., van Rensburg, M.J.J., Bray, J.E., Earle, S.G., Ford, S.A., Jolley, K.A. and McCarthy, N.D., 2013. MLST revisited: the gene-by-gene approach to bacterial genomics. *Nature Reviews Microbiology*, 11(10), pp.728–736.

Man, S.M., 2011. The clinical importance of emerging Campylobacter species. *Nature Reviews Gastroenterology & Hepatology*, 8(12), pp.669–685.

Man, S.M., Zhang, L., Day, A.S., Leach, S.T., Lemberg, D.A. and Mitchell, H., 2010. Campylobacter concisus and other Campylobacter species in children with newly diagnosed Crohn's disease. *Inflammatory Bowel Diseases*, 16(6), pp.1008–1016.

Mansfield, L.S., Bell, J.A., Wilson, D.L., Murphy, A.J., Elsheikha, H.M., Rathinam, V.A.K., Fierro, B.R., Linz, J.E. and Young, V.B., 2007. C57BL/6 and Congenic Interleukin-10-Deficient Mice Can Serve as Models of Campylobacter jejuni Colonization and Enteritis. *Infection and Immunity*, 75(3), pp.1099–1115.

Marri, P.R., Hao, W. and Golding, G.B., 2006. Gene Gain and Gene Loss in Streptococcus: Is It Driven by Habitat? *Molecular Biology and Evolution*, 23(12), pp.2379–2391.

Martin, P.M.V., Mathiot, J., Ipero, J., Georges, A.J. and Georges-Courbot, M.C., 1988. Antibody response to Campylobacter coli in children during intestinal infection and carriage. *Journal of Clinical Microbiology*, 26(7), pp.1421–1424.

Mazel, D. and Davies, J., 1999. Antibiotic resistance in microbes. *Cellular and Molecular Life Sciences (CMLS)*, 56(9–10), pp.742–754.

McFadyean, F. and Stockman, S., 1913. *Report of the Departmental Committee appointed by the Board of Agriculture and Fisheries to inquire into epizootic abortion. Appendix to part III, Abortion in sheep* [Online]. London: HMSO by Eyre and Spottiswoode.

McInerney, J.O., McNally, A. and O'Connell, M.J., 2017. Why prokaryotes have pangenomes. *Nature Microbiology*, 2(March), pp.1–5.

McKenna, M., 2013. Antibiotic resistance: The last resort. *Nature*, 499(7459), pp.394–396.

McMichael, A.J., 2004. Environmental and social influences on emerging infectious diseases: past, present and future. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359(1447), pp.1049–1058.

McNally, A., Cheng, L., Harris, S.R. and Corander, J., 2013. The Evolutionary Path

to Extraintestinal Pathogenic, Drug-Resistant Escherichia coli Is Marked by Drastic Reduction in Detectable Recombination within the Core Genome. *Genome Biology and Evolution*, 5(4), pp.699–710.

McSweegan, E. and Walker, R.I., 1986. Identification and characterization of two Campylobacter jejuni adhesins for cellular and mucous substrates. *Infection and Immunity*, 53(1), pp.141–148.

Méric, G., Mageiros, L., Pensar, J., Laabei, M., Yahara, K., Pascoe, B., Kittiwan, N., Tadee, P., Post, V., Lamble, S., Bowden, R., Bray, J.E., Morgenstern, M., Jolley, K.A., Maiden, M.C.J., Feil, E.J., Didelot, X., Miragaia, M., de Lencastre, H., Moriarty, T.F., Rohde, H., Massey, R., Mack, D., Corander, J. and Sheppard, S.K., 2018. Disease-associated genotypes of the commensal skin bacterium Staphylococcus epidermidis. *Nature Communications*, 9(1), p.5034.

Milkman, R., 1973. Electrophoretic Variation in Escherichia coli from Natural Sources. *Science*, 182(4116), pp.1024–1026.

Mira, A., Martín-Cuadrado, A.B., D'Auria, G. and Rodríguez-Valera, F., 2010. The bacterial pan-genome:a new paradigm in microbiology. *International microbiology : the official journal of the Spanish Society for Microbiology*, 13(2), pp.45–57.

Monteville, M.R., 2003. Maximal adherence and invasion of INT 407 cells by Campylobacter jejuni requires the CadF outer-membrane protein and microfilament reorganization. *Microbiology*, 149(1), pp.153–165.

Moran, L., Kelly, C. and Madden, R.H., 2009. Factors affecting the recovery of Campylobacter spp. from retail packs of raw, fresh chicken using ISO 10272-1:2006. *Letters in Applied Microbiology*, 48(5), pp.628–632.

Moran, N.A. and Wernegreen, J.J., 2000. Lifestyle evolution in symbiotic bacteria: insights from genomics. *Trends in Ecology & Evolution*, 15(8), pp.321–326.

Morley, L., McNally, A., Paszkiewicz, K., Corander, J., Méric, G., Sheppard, S.K., Blom, J. and Manning, G., 2015. Gene Loss and Lineage-Specific Restriction-Modification Systems Associated with Niche Differentiation in the Campylobacter jejuni Sequence Type 403 Clonal Complex. *Applied and Environmental Microbiology*, 81(11), pp.3641–3647.

Mourkas, E., Florez-Cuadrado, D., Pascoe, B., Calland, J.K., Bayliss, S.C., Mageiros, L., Méric, G., Hitchings, M.D., Quesada, A., Porrero, C., Ugarte-Ruiz, M., Gutiérrez-Fernández, J., Domínguez, L. and Sheppard, S.K., 2019. Gene pool transmission of multidrug resistance among Campylobacter from

livestock, sewage and human disease. *Environmental Microbiology*, 21(12), pp.4597–4613.

Mourkas, E., Taylor, A.J., Méric, G., Bayliss, S.C., Pascoe, B., Mageiros, L., Calland, J.K., Hitchings, M.D., Ridley, A., Vidal, A., Forbes, K.J., Strachan, N.J.C., Parker, C.T., Parkhill, J., Jolley, K.A., Cody, A.J., Maiden, M.C.J., Kelly, D.J. and Sheppard, S.K., 2020. Agricultural intensification and the evolution of host specialism in the enteric pathogen Campylobacter jejuni. *Proceedings of the National Academy of Sciences*, 117(20), pp.11018–11028.

Mughini-Gras, L., Smid, J.H., Wagenaar, J.A., DE Boer, A., Havelaar, A.H., Friesema, I.H.M., French, N.P., Graziani, C., Busani, L. and Van Pelt, W., 2014. Campylobacteriosis in returning travellers and potential secondary transmission of exotic strains. *Epidemiology and infection*, 142(6), pp.1277–88.

Murray, S., Pascoe, B., Méric, G., Mageiros, L., Yahara, K., Hitchings, M.D., Friedmann, Y., Wilkinson, T.S., Gormley, F.J., Mack, D., Bray, J.E., Lamble, S., Bowden, R., Jolley, K.A., Maiden, M.C.J., Wendlandt, S., Schwarz, S., Corander, J., Fitzgerald, J.R. and Sheppard, S.K., 2017. Recombination-Mediated Host Adaptation by Avian Staphylococcus aureus. *Genome Biology and Evolution*, 9(4), pp.830–842.

Muyzer, G., de Waal, E.C. and Uitterlinden, A.G., 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied and Environmental Microbiology*, 59(3), pp.695–700.

Nachamkin, I., 2002. Chronic effects of Campylobacter infection. *Microbes and Infection*, 4(4), pp.399–403.

Nelson, M.I. and Holmes, E.C., 2007. The evolution of epidemic influenza. *Nature Reviews Genetics*, 8(3), pp.196–205.

Nielsen, L.N., Sheppard, S.K., McCarthy, N.D., Maiden, M.C.J., Ingmer, H. and Krogfelt, K.A., 2010. MLST clustering of Campylobacter jejuni isolates from patients with gastroenteritis, reactive arthritis and Guillain-Barré syndrome. *Journal of applied microbiology*, 108(2), pp.591–9.

Noto, M.J., Kreiswirth, B.N., Monk, A.B. and Archer, G.L., 2008. Gene Acquisition at the Insertion Site for SCCmec, the Genomic Island Conferring Methicillin Resistance in Staphylococcus aureus. *Journal of Bacteriology*, 190(4), pp.1276–1283.

Nsofor, C.A., 2016. Pulsed-Field Gel Electrophoresis (PFGE): Principles and Applications in Molecular Epidemiology: A Review. *Int. J. Curr. Res. Med. Sci*, 2(2), pp.38–51.

Ogden, I.D., Dallas, J.F., MacRae, M., Rotariu, O., Reay, K.W., Leitch, M., Thomson, A.P., Sheppard, S.K., Maiden, M., Forbes, K.J. and Strachan, N.J.C., 2009. Campylobacter Excreted into the Environment by Animal Sources: Prevalence, Concentration Shed, and Host Association. *Foodborne Pathogens and Disease*, 6(10), pp.1161–1170.

On, S.L.W., 2013. Isolation, identification and subtyping of Campylobacter: Where to from here? *Journal of Microbiological Methods*, 95(1), pp.3–7.

Otto, S.P. and Lenormand, T., 2002. Resolving the paradox of sex and recombination. *Nature Reviews Genetics*, 3(4), pp.252–261.

Oyarzabal, O.A., Macklin, K.S., Barbaree, J.M. and Miller, R.S., 2005. Evaluation of Agar Plates for Direct Enumeration of Campylobacter spp. from Poultry Carcass Rinses. *Applied and Environmental Microbiology*, 71(6), pp.3351–3354.

Parkhill, J., Wren, B.W., Mungall, K., Ketley, J.M., Churcher, C., Basham, D., Chillingworth, T., Davies, R.M., Feltwell, T., Holroyd, S., Jagels, K., Karlyshev, A. V., Moule, S., Pallen, M.J., Penn, C.W., Quail, M.A., Rajandream, M.-A., Rutherford, K.M., van Vliet, A.H.M., Whitehead, S. and Barrell, B.G., 2000. The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences. *Nature*, 403(6770), pp.665–668.

Pascoe, B., Méric, G., Murray, S., Yahara, K., Mageiros, L., Bowen, R., Jones, N.H., Jeeves, R.E., Lappin-Scott, H.M., Asakura, H. and Sheppard, S.K., 2015. Enhanced biofilm formation and multi-host transmission evolve from divergent genetic backgrounds in C ampylobacter jejuni. *Environmental Microbiology*, 17(11), pp.4779–4789.

Pascoe, B., Méric, G., Yahara, K., Wimalarathna, H., Murray, S., Hitchings, M.D., Sproston, E.L., Carrillo, C.D., Taboada, E.N., Cooper, K.K., Huynh, S., Cody, A.J., Jolley, K.A., Maiden, M.C.J., McCarthy, N.D., Didelot, X., Parker, C.T. and Sheppard, S.K., 2017. Local genes for local bacteria: Evidence of allopatry in the genomes of transatlantic Campylobacter populations. *Molecular Ecology*, 26(17), pp.4497–4508.

Pascoe, B., Schiaffino, F., Murray, S., Méric, G., Bayliss, S.C., Hitchings, M.D., Mourkas, E., Calland, J.K., Burga, R., Penataro Yori, P., Jolley, K.A., Cooper,

K.K., Parker, C.T., Paredes Olortegui, M., Kosek, M.N. and Sheppard, S.K., 2020. Genomic epidemiology of Campylobacter jejuni associated with asymptomatic pediatric infection in the Peruvian Amazon. *bioRxiV*.

Penner, J.L. and Hennessy, J.N., 1980. Passive hemagglutination technique for serotyping Campylobacter fetus subsp. jejuni on the basis of soluble heat-stable antigens. *Journal of Clinical Microbiology*, 12(6), pp.732–737.

Petes, T.D., 2001. Meiotic recombination hot spots and cold spots. *Nature Reviews Genetics*, 2(5), pp.360–369.

Pittman, M.S., Elvers, K.T., Lee, L., Jones, M.A., Poole, R.K., Park, S.F. and Kelly, D.J., 2007. Growth of Campylobacter jejuni on nitrate and nitrite: electron transport to NapA and NrfA via NrfH and distinct roles for NrfA and the globin Cgb in protection against nitrosative stress. *Molecular Microbiology*, 63(2), pp.575–590.

Poirel, L., Dortet, L., Bernabeu, S. and Nordmann, P., 2011. Genetic Features of bla NDM-1 -Positive Enterobacteriaceae. *Antimicrobial Agents and Chemotherapy*, 55(11), pp.5403–5407.

Price, M.N., Dehal, P.S. and Arkin, A.P., 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE*, 5(3), p.e9490.

Qin, S., Wang, Y., Zhang, Q., Chen, X., Shen, Z., Deng, F., Wu, C. and Shen, J., 2012. Identification of a Novel Genomic Island Conferring Resistance to Multiple Aminoglycoside Antibiotics in Campylobacter coli. *Antimicrobial Agents and Chemotherapy*, 56(10), pp.5332–5339.

Qin, S., Wang, Y., Zhang, Q., Zhang, M., Deng, F., Shen, Z., Wu, C., Wang, S., Zhang, J. and Shen, J., 2014. Report of ribosomal RNA methylase gene erm(B) in multidrug-resistant Campylobacter coli. *Journal of Antimicrobial Chemotherapy*, 69(4), pp.964–968.

Rambaut, A., Lam, T.T., Max Carvalho, L. and Pybus, O.G., 2016. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evolution*, 2(1), p.vew007.

Raoult, D., 2014. Editorial: Emerging clones of bacterial epidemics in the genomic area. *Clinical Microbiology and Infection*, 20(5), pp.371–372.

Read, T.D. and Massey, R.C., 2014. Characterizing the genetic basis of bacterial phenotypes using genome-wide association studies: a new direction for bacteriology. *Genome Medicine*, 6(11), p.109.

Richardson, E.J., Bacigalupe, R., Harrison, E.M., Weinert, L.A., Lycett, S., Vrieling,

M., Robb, K., Hoskisson, P.A., Holden, M.T.G., Feil, E.J., Paterson, G.K., Tong, S.Y.C., Shittu, A., van Wamel, W., Aanensen, D.M., Parkhill, J., Peacock, S.J., Corander, J., Holmes, M. and Fitzgerald, J.R., 2018. Gene exchange drives the ecological success of a multi-host bacterial pathogen. *Nature Ecology & Evolution*, 2(9), pp.1468–1478.

Rieux, A. and Balloux, F., 2016. Inferences from tip-calibrated phylogenies: a review and a practical guide. *Molecular Ecology*, 25(9), pp.1911–1924.

Roberts, R.B., de Lencastre, A., Eisner, W., Severina, E.P., Shopsin, B., Kreiswirth, B.N. and Tomasz, A., 1998. Molecular Epidemiology of Methicillin-Resistant Staphylococcus aureus in 12 New York Hospitals. *Journal of Infectious Diseases*, 178(1), pp.164–171.

Rodríguez-Beltrán, J., Tourret, J., Tenaillon, O., López, E., Bourdelier, E., Costas, C., Matic, I., Denamur, E. and Blázquez, J., 2015. High Recombinant Frequency in Extraintestinal Pathogenic Escherichia coli Strains. *Molecular Biology and Evolution*, 32(7), pp.1708–1716.

Rogol, M., Sechter, I., Falk, H., Shtark, Y., Alfi, S., Greenberg, Z. and Mizrachi, R., 1983. Waterborne outbreak of Campylobacter enteritis. *European Journal of Clinical Microbiology*, 2(6), pp.588–590.

Ryall, B., Eydallin, G. and Ferenci, T., 2012. Culture History and Population Heterogeneity as Determinants of Bacterial Adaptation: the Adaptomics of a Single Environmental Transition. *Microbiology and Molecular Biology Reviews*, 76(3), pp.597–625.

Saber, M.M. and Shapiro, B.J., 2020. Benchmarking bacterial genome-wide association study methods using simulated genomes and phenotypes. *Microbial Genomics*, 6(3).

Saitou, N. and Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4(4), pp.406–425.

Salyers, A.A., Gupta, A. and Wang, Y., 2004. Human intestinal bacteria as reservoirs for antibiotic resistance genes. *Trends in microbiology*, 12(9), pp.412–6.

Schierup, M.H. and Hein, J., 2000. Recombination and the Molecular Clock. *Molecular Biology and Evolution*, 17(10), pp.1578–1579.

Schubert, S., Darlu, P., Clermont, O., Wieser, A., Magistro, G., Hoffmann, C., Weinert, K., Tenaillon, O., Matic, I. and Denamur, E., 2009. Role of

Intraspecies Recombination in the Spread of Pathogenicity Islands within the Escherichia coli Species. *PLoS Pathogens*, 5(1), p.e1000257.

Schwarz, S., Kehrenberg, C. and Walsh, T.R., 2001. Use of antimicrobial agents in veterinary medicine and food animal production. *International Journal of Antimicrobial Agents*, 17(6), pp.431–437.

Schwille-Kiuntke, J., Enck, P., Zendler, C., Krieg, M., Polster, A. V., Klosterhalfen, S., Autenrieth, I.B., Zipfel, S. and Frick, J.-S., 2011. Postinfectious irritable bowel syndrome: follow-up of a patient cohort of confirmed cases of bacterial infection with Salmonella or Campylobacter. *Neurogastroenterology & Motility*, 23(11), pp.e479–e488.

Sebald, M. and Veron, M., 1963. Teneur en bases de L'ADN et classification de vibrions. *Annales de l'Institut Pasteur* [Online], 105, pp.897–910.

Selander, R.K., Caugant, D.A., Ochman, H., Musser, J.M., Gilmour, M.N. and Whittam, T.S., 1986. Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. *Applied and Environmental Microbiology*, 51(5), pp.873–884.

Shao, D., Shi, Z., Wei, J. and Ma, Z., 2011. A brief review of foodborne zoonoses in China. *Epidemiology and infection*, 139(10), pp.1497–504.

Shapiro, B.J., 2016. How clonal are bacteria over time? *Current Opinion in Microbiology*, 31, pp.116–123.

Shapiro, B.J., 2017. The population genetics of pangenomes. *Nature Microbiology*, 2(12), pp.1574–1574.

Sharp, P.M. and Hahn, B.H., 2011. Origins of HIV and the AIDS Pandemic. *Cold Spring Harbor Perspectives in Medicine*, 1(1), pp.a006841–a006841.

Sheppard, S.K., Cheng, L., Méric, G., de Haan, C.P.A., Llarena, A.-K., Marttinen, P., Vidal, A., Ridley, A., Clifton-Hadley, F., Connor, T.R., Strachan, N.J.C., Forbes, K., Colles, F.M., Jolley, K.A., Bentley, S.D., Maiden, M.C.J., Hänninen, M.-L., Parkhill, J., Hanage, W.P. and Corander, J., 2014. Cryptic ecology among host generalist Campylobacter jejuni in domestic animals. *Molecular Ecology*, 23(10), pp.2442–2451.

Sheppard, S.K., Colles, F.M., McCarthy, N.D., Strachan, N.J.C., Ogden, I.D., Forbes, K.J., Dallas, J.F. and Maiden, M.C.J., 2011. Niche segregation and genetic structure of Campylobacter jejuni populations from wild and agricultural host species. *Molecular ecology*, 20(16), pp.3484–90.

Sheppard, Samuel K., Dallas, J.F., MacRae, M., McCarthy, N.D., Sproston, E.L.,

Gormley, F.J., Strachan, N.J.C., Ogden, I.D., Maiden, M.C.J. and Forbes, K.J., 2009a. Campylobacter genotypes from food animals, environmental sources and clinical disease in Scotland 2005/6. *International Journal of Food Microbiology*, 134(1–2), pp.96–103.

Sheppard, Samuel K., Dallas, J.F., Strachan, N.J.C., MacRae, M., McCarthy, N.D., Wilson, D.J., Gormley, F.J., Falush, D., Ogden, I.D., Maiden, M.C.J. and Forbes, K.J., 2009b. Campylobacter genotyping to determine the source of human infection. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 48(8), pp.1072–8.

Sheppard, S.K., Dallas, J.F., Wilson, D.J., Strachan, N.J.C., McCarthy, N.D., Jolley, K.A., Colles, F.M., Rotariu, O., Ogden, I.D., Forbes, K.J. and Maiden, M.C.J., 2010. Evolution of an Agriculture-Associated Disease Causing Campylobacter coli Clade: Evidence from National Surveillance Data in Scotland. *PLoS ONE*, 5(12), p.e15708.

Sheppard, S.K., Didelot, X., Meric, G., Torralbo, A., Jolley, K.A., Kelly, D.J., Bentley, S.D., Maiden, M.C.J., Parkhill, J. and Falush, D., 2013. Genome-wide association study identifies vitamin B5 biosynthesis as a host specificity factor in Campylobacter. *Proceedings of the National Academy of Sciences of the United States of America*, 110(29), pp.11923–7.

Sheppard, S.K., Guttman, D.S. and Fitzgerald, J.R., 2018. Population genomics of bacterial host adaptation. *Nature Reviews Genetics*, 19(9), pp.549–565.

Sheppard, S.K., Jolley, K.A. and Maiden, M.C.J., 2012. A Gene-By-Gene Approach to Bacterial Population Genomics: Whole Genome MLST of Campylobacter. *Genes*, 3(2), pp.261–277.

Skirrow, M.B., 1977. Campylobacter enteritis: a ‘new’ disease. *BMJ*, 2(6078), pp.9–11.

Smith, A.B., Pisani, D., Mackenzie-Dodds, J.A., Stockley, B., Webster, B.L. and Littlewood, D.T.J., 2006. Testing the Molecular Clock: Molecular and Paleontological Estimates of Divergence Times in the Echinoidea (Echinodermata). *Molecular Biology and Evolution*, 23(10), pp.1832–1851.

Smith, J.M., Smith, N.H., O’Rourke, M. and Spratt, B.G., 1993. How clonal are bacteria? *Proceedings of the National Academy of Sciences*, 90(10), pp.4384–4388.

Smith, J.T., Amador, S., McGonagle, C.J., Needle, D., Gibson, R. and Andam, C.P., 2020. Population genomics of Staphylococcus pseudintermedius in

companion animals in the United States. *Communications Biology*, 3(1), p.282.

Snelling, W.J., McKenna, J.P., Hack, C.J., Moore, J.E. and Dooley, J.S.G., 2006. An examination of the diversity of a novel Campylobacter reservoir. *Archives of Microbiology*, 186(1), pp.31–40.

Solomon, J.M. and Grossman, A.D., 1996. Who's competent and when: regulation of natural genetic competence in bacteria. *Trends in Genetics*, 12(4), pp.150–155.

Sproston, E.L., Ogden, I.D., MacRae, M., Dallas, J.F., Sheppard, S.K., Cody, A.J., Colles, F.M., Wilson, M.J., Forbes, K.J. and Strachan, N.J.C., 2011. Temporal Variation and Host Association in the Campylobacter Population in a Longitudinal Ruminant Farm Study. *Applied and Environmental Microbiology*, 77(18), pp.6579–6586.

Sproston, E.L., Wimalarathna, H.M.L. and Sheppard, S.K., 2018. Trends in fluoroquinolone resistance in Campylobacter. *Microbial Genomics*, 4(8), pp.1–8.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), pp.1312–1313.

Suerbaum, S., Smith, J.M., Bapumia, K., Morelli, G., Smith, N.H., Kunstmann, E., Dyrek, I. and Achtman, M., 1998. Free recombination within Helicobacter pylori. *Proceedings of the National Academy of Sciences*, 95(21), pp.12619–12624.

Swaminathan, A., Torresi, J., Schlagenhauf, P., Thursky, K., Wilder-Smith, A., Connor, B.A., Schwartz, E., VonSonnenberg, F., Keystone, J. and O'Brien, D.P., 2009. A global study of pathogens and host risk factors associated with infectious gastrointestinal disease in returned international travellers. *Journal of Infection*, 59(1), pp.19–27.

Taylor, D.E., 1986. Plasmid-mediated tetracycline resistance in Campylobacter jejuni: expression in Escherichia coli and identification of homology with streptococcal class M determinant. *Journal of Bacteriology*, 165(3), pp.1037–1039.

Taylor, D.E., Garner, R.S. and Allan, B.J., 1983. Characterization of tetracycline resistance plasmids from Campylobacter jejuni and Campylobacter coli. *Antimicrobial Agents and Chemotherapy*, 24(6), pp.930–935.

Teuber, M., 2001. Veterinary use and antibiotic resistance. *Current Opinion in*

Microbiology, 4(5), pp.493–499.

Thépault, A., Méric, G., Rivoal, K., Pascoe, B., Mageiros, L., Touzain, F., Rose, V., Béven, V., Chemaly, M. and Sheppard, S.K., 2017. Genome-Wide Identification of Host-Segregating Epidemiological Markers for Source Attribution in Campylobacter jejuni. *Applied and Environmental Microbiology*, 83(7), pp.1–13.

Thomas, C.M. and Nielsen, K.M., 2005. Mechanisms of, and Barriers to, Horizontal Gene Transfer between Bacteria. *Nature Reviews Microbiology*, 3(9), pp.711–721.

Thorns, C.J., 2000. Bacterial food-borne zoonoses Changing perspectives : 20th and 21 st Centuries. *Revi. Sci. Tech. Off. Int. Epiz.* [Online], 19(1), pp.226–239.

To, T.-H., Jung, M., Lycett, S. and Gascuel, O., 2016. Fast Dating Using Least-Squares Criteria and Algorithms. *Systematic Biology*, 65(1), pp.82–97.

Toft, C. and Andersson, S.G.E., 2010. Evolutionary microbial genomics: insights into bacterial host adaptation. *Nature Reviews Genetics*, 11(7), pp.465–475.

Viana, D., Blanco, J., Tormo-Más, M.Á., Selva, L., Guinane, C.M., Baselga, R., Corpa, J.M., Lasa, Í., Novick, R.P., Fitzgerald, J.R. and Penadés, J.R., 2010. Adaptation of Staphylococcus aureus to ruminant and equine hosts involves SaPI-carried variants of von Willebrand factor-binding protein. *Molecular Microbiology*, 77(6), pp.1583–1594.

Vos, M., 2009. Why do bacteria engage in homologous recombination? *Trends in Microbiology*, 17(6), pp.226–232.

Vos, M. and Didelot, X., 2009. A comparison of homologous recombination rates in bacteria and archaea. *The ISME Journal*, 3(2), pp.199–208.

Wagenaar, J.A., van Bergen, M.A.P., Blaser, M.J., Tauxe, R. V, Newell, D.G. and van Putten, J.P.M., 2014. Campylobacter fetus Infections in Humans: Exposure and Disease. *Clinical Infectious Diseases*, 58(11), pp.1579–1586.

Wardeh, M., Risley, C., McIntyre, M.K., Setzkorn, C. and Baylis, M., 2015. Database of host-pathogen and related species interactions, and their global distribution. *Scientific Data*, 2(1), p.150049.

Wassenaar, T.M., Fry, B.N. and van der Zeijst, B.A.M., 1995. Variation of the flagellin gene locus of Campylobacter jejuni by recombination and horizontal gene transfer. *Microbiology*, 141(1), pp.95–101.

Weinert, L.A., Chaudhuri, R.R., Wang, J., Peters, S.E., Corander, J., Jombart, T.,

Baig, A., Howell, K.J., Vehkala, M., Välimäki, N., Harris, D., Chieu, T.T.B., Van Vinh Chau, N., Campbell, J., Schultsz, C., Parkhill, J., Bentley, S.D., Langford, P.R., Rycroft, A.N., Wren, B.W., Farrar, J., Baker, S., Hoa, N.T., Holden, M.T.G., Tucker, A.W. and Maskell, D.J., 2015. Genomic signatures of human and animal disease in the zoonotic pathogen Streptococcus suis. *Nature Communications*, 6(1), p.6740.

Weinert, L.A., Welch, J.J., Suchard, M.A., Lemey, P., Rambaut, A. and Fitzgerald, J.R., 2012. Molecular dating of human-to-bovid host jumps by Staphylococcus aureus reveals an association with the spread of domestication. *Biology Letters*, 8(5), pp.829–832.

Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., Klemm, A., Flicek, P., Manolio, T., Hindorff, L. and Parkinson, H., 2014. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Research*, 42(D1), pp.D1001–D1006.

Werner, G., Hildebrandt, B. and Witte, W., 2003. Linkage of erm (B) and aadE-sat4-aphA-3 in Multiple-Resistant Enterococcus faecium Isolates of Different Ecological Origins. *Microbial Drug Resistance*, 9(supplement 1), pp.9–16.

Whitman, W.B., Coleman, D.C. and Wiebe, W.J., 1998. Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences*, 95(12), pp.6578–6583.

WHO, 2017. *WHO publishes list of bacteria for which new antibiotics are urgently needed*.

Wieser, A., Schneider, L., Jung, J. and Schubert, S., 2012. MALDI-TOF MS in microbiological diagnostics—identification of microorganisms and beyond (mini review). *Applied Microbiology and Biotechnology*, 93(3), pp.965–974.

Wilson, D.J., Gabriel, E., Leatherbarrow, A.J.H., Cheesbrough, J., Gee, S., Bolton, E., Fox, A., Hart, C.A., Diggle, P.J. and Fearnhead, P., 2009. Rapid Evolution and the Importance of Recombination to the Gastroenteric Pathogen Campylobacter jejuni. *Molecular Biology and Evolution*, 26(2), pp.385–397.

Woolhouse, M.E.J., 2001. Population Biology of Multihost Pathogens. *Science*, 292(5519), pp.1109–1112.

Woolhouse, M.E.J., Haydon, D.T. and Antia, R., 2005. Emerging pathogens: the epidemiology and evolution of species jumps. *Trends in Ecology & Evolution*, 20(5), pp.238–244.

Worobey, M., Han, G.-Z. and Rambaut, A., 2014. A synchronized global sweep of

the internal genes of modern avian influenza virus. *Nature*, 508(7495), pp.254–257.

Yahara, K., Didelot, X., Jolley, K.A., Kobayashi, I., Maiden, M.C.J., Sheppard, S.K. and Falush, D., 2016. The Landscape of Realized Homologous Recombination in Pathogenic Bacteria. *Molecular Biology and Evolution*, 33(2), pp.456–471.

Young, K.T., Davis, L.M. and DiRita, V.J., 2007. Campylobacter jejuni: molecular biology and pathogenesis. *Nature Reviews Microbiology*, 5(9), pp.665–679.

Zhang, L., Man, S.M., Day, A.S., Leach, S.T., Lemberg, D.A., Dutt, S., Stormon, M., Otley, A., O'Loughlin, E. V., Magoffin, A., Ng, P.H.Y. and Mitchell, H., 2009. Detection and Isolation of Campylobacter Species Other than C. jejuni from Children with Crohn's Disease. *Journal of Clinical Microbiology*, 47(2), pp.453–455.